

A MATHEMATICAL ANALYSIS OF THE EFFECTS OF
COMPONENT FAILURES ON SYSTEM PERFORMANCE

By

Richard Charles Dubes

AN ABSTRACT OF A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering

1962

ABSTRACT

A MATHEMATICAL ANALYSIS OF THE EFFECTS OF COMPONENT FAILURES ON SYSTEM PERFORMANCE

by Richard Charles Dubes

The requirements of modern technology necessitate extremely large and complex systems. As the number of components increases, so does the chance that the system's operational worth will be degraded because of component failures. This thesis proposes a mathematical model, called the General Model, which analyzes the effects of such failures on a system's operational worth, as measured by reliability and utility functions. This model differs from existing models in that it includes the effects of both drift and catastrophic failures on the performance of redundant systems.

The General Model separates the analysis into a set of drift problems by defining states for a system. A system changes states when a component fails catastrophically. Since a system's drift properties change when a component fails, a drift analysis is made for each state.

A system is classified as either loaded or non-loaded, depending on whether or not the catastrophic failure of a redundant component changes the failure rates of other components. Techniques for finding the probability distribution of the system's

Richard Charles Dubes

states are presented for both cases. The components can fail catastrophically in either one or two modes.

Two algorithms are presented for finding reliability expressions when a system is non-loaded and drift is neglected. These algorithms are applicable to any active redundant system for any component failure distribution. The first algorithm applies when components have one failure mode; the second, when two failure modes are allowed. A procedure for simplifying the resulting reliability expressions is also presented.

Two types of component drift are discussed. In the first, the amount of drift depends on the length of time the component operates; in the second, the drift is a function of the states the system passes through and how long it remains in each state. Techniques for determining these drift properties and for relating them to system drift properties are provided.

A MATHEMATICAL ANALYSIS OF THE EFFECTS OF
COMPONENT FAILURES ON SYSTEM PERFORMANCE

By

Richard Charles Dubes

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering

1962

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Dr. H. E. Koenig for his guidance in the preparation of this thesis. He also wishes to thank Dr. J. H. Stapleton, Dr. M. B. Reed, and Dr. G. P. Weeg for the fruitful discussions concerning this thesis.

CONTENTS

	Page
ACKNOWLEDGMENTS	ii
LIST OF ILLUSTRATIONS	iv
LIST OF APPENDICES.	v
Chapter	
I. INTRODUCTION.	1
Component Failures	
Classification of Systems	
Neutralizing Failures	
II. MATHEMATICAL MODELS OF FAILURE ANALYSES	8
Foundations of the General Model	
Existing Models	
III. CATASTROPHIC FAILURES	27
One Catastrophic Failure Mode	
Two Catastrophic Failure Modes	
IV. DRIFT EFFECTS	70
The Drift Problem	
Distributions of System Parameters	
at One Time	
Time Variations	
Conclusions	
LIST OF REFERENCES.	123

LIST OF ILLUSTRATIONS

Figure	Page
1. Example of a State Table and Corresponding Reliability Diagram	31
2. Examples of Reliability Diagrams.	32
3. Reliability Diagrams.	33
4. Example of a State Table, Transition Chart and Reliability Diagram	38
5. Example of a Reliability Diagram and State Table	40
6. Markov Transition Chart Corresponding to Fig. 5	44
7. Partial Reliability Table for Fig. 3A	53
8. Tabulations of a Simplification Procedure for Fig. 7.	54
9. Example of a State Table and Corresponding Reliability Diagram	60
10. Example of a Reliability Diagram.	67
11. Example of Step 1 of the Two-Failure Algorithm	68
12. Example of Step 2 of the Two-Failure Algorithm	69
13. Histogram of Measured Data.	85
14. Power Supply System and Corresponding State Table and Transition Diagram.	87

LIST OF APPENDICES

Appendix	Page
A. Some Basic Definitions from Probability Theory	98
B. Continuous Parameter Markov Chains with a Finite Number of States.	101
C. The Application of Markov Chains to the Analysis of Loaded Systems	110
D. Justification of the Two-Failure Algorithm.	115
E. Derivation of an Approximation to the Mean and Variance of a Function of Random Variables . . .	120

CHAPTER I

INTRODUCTION

The technology that sends man into space and propels missiles from continent to continent has an insatiable appetite for more complex and more sophisticated systems¹. National security necessitates the gains in operational worth that result from larger systems. However, these gains are offset by an increase in the chance of system failure, a consequence of using more components. The need for more reliable systems requires no further elaboration.

Any scheme that is directed at decreasing the chance of system failure is, by common usage, termed a reliability program². An interesting history of the efforts in this direction has been presented by Ryerson (6). As discussed by Chorafas (7 Chapter 40), a successful reliability program begins with specifications and proceeds through design, development, manufacture, and field testing. Component failure data is fed back from each stage in a continual re-design cycle.

¹For example, Koenig (1) has stated that the SAGE continental defense system contains over 1,252,000 individual tubes, diodes, resistors, capacitors, inductors, and printed circuits. Perry (2) reports that the NIKE-HERCULES missile system consists of about 1,500,000 individual parts.

²A wide variety of reliability programs have been proposed in the literature, e.g., by Saltz (3), Greene (4), and Dreste (5).

A most important step in any reliability program is the failure analysis which predicts, on the basis of component failure data¹, whether the system meets its reliability objectives. This thesis presents a mathematical model for accomplishing such a failure analysis analytically.

For purposes of a failure analysis, a system is presumed to have been optimized on the basis of criteria prescribed by specifications, such as the mean-squared error for a control system or the signal-to-noise ratio for a transmitter. Suitable measures of these performance criteria are assumed to be defined so that a system's operational worth can be judged quantitatively. Sufficient data on the failure tendencies of the components are also assumed to be available. The failure analysis presented in this thesis uses these data, along with information on component connections, to determine the measures of operational worth.

Component failures

Component failures are ordinarily separated into two classes (e.g., see Feyerherm (9)) called catastrophic, or chance, and drift, or degradation, failures. The distinguishing features of these classes are discussed in this section.

Connor (10) describes catastrophic failure as follows.
"Chance (catastrophic) failure will include those physical

¹The precision of any prediction of system reliability rests on the amount and accuracy of component failure data. Statistical sampling plans for obtaining this data will not be presented in this thesis. An excellent exposition of such plans has been given by Epstein (8).

occurrences (without a central tendency in time) which are beyond the engineering resolution of normal techniques for product design, manufacture and operation, and cause cessation of performance requiring corrective maintenance.

When a component fails catastrophically, its characteristics change suddenly, a discontinuity in operation occurs, rendering the component operationally useless. The term "failure" usually refers to catastrophic failure which is the only type of failure considered in the majority of books and papers on reliability.

A component's chance of failing catastrophically is determined by two factors--the component's inherent strength and the environment. The former is built into the component by the manufacturing process. Uncontrollable variations in raw materials and fabrication techniques cause variations in the susceptibility of presumably identical components to failure. This inherent strength can change with time as the component deteriorates with age. The environment is the totality of all random and deterministic factors in the operating medium which can cause changes in the characteristics of a component. If failure data is to be used to predict failure tendencies, it must be taken under the same environment that will be experienced in actual use.

In contrast to the spontaneity of catastrophic failure, drift is a gradual transition to less desirable component properties. The point at which a component fails because of drift depends on the sensitivity of system performance to that component's properties, on the drift characteristics of other components, and on the amount of drift in system performance that

can be tolerated. The drift failure of a system, which is determined in terms of prescribed tolerance limits, can be caused by a relatively slight drift in a group of components in the same direction or by extreme drift in a single component. Also, two components might drift in opposite directions so as to compensate for one another. For these reasons, the drift failure of a component cannot be uniquely related to that for a system. Deleterious drift in one situation can be tolerable in another.

Classification of systems

The mathematical techniques which are applied to a failure analysis depend to a great extent on the manner in which the system is used. Systems are classified, by usage, as intermittent or continuous and as repairable or non-repairable.

A continuous system is one which is turned on and allowed to operate until it fails and in which components can fail catastrophically at any time. That is, the failure-causing environment is present whenever the system is operating. Examples of continuous systems are airplanes and missiles during flight, the control systems driving a scanning radar antenna, and the generators supplying a power distribution system.

An intermittent system, on the other hand, is one whose periods of use and non-use alternate, usually in a random fashion. Such systems are often subsystems which operate at the command of a larger system. Examples are the braking system in an automobile and the arithmetic units in a digital computer. An excellent failure analysis of intermittent systems has been presented by Flehinger (11).

In a repairable system, components are replaced when they fail or according to a prescribed maintenance policy. The time needed to locate, replace, and repair the failed component is included in the failure study. In a non-repairable system, each component operates until it or the system fails; both component and system failures are permanent.

In this thesis, all systems are assumed to be continuous and non-repairable.

Neutralizing failures

The severity of component drift can often be reduced by proper design, especially in electronic circuits. The two standard design procedures for effecting this reduction are called the "worst case" method, discussed by Ashcraft and Hochwald (12) and Suran (13) and "marginal checking", examined by Chorafas (7 p. 370), Patterson (14), Brown and Dennis (15), and Drenick (16). In both methods, those component parameters to which the system is most sensitive are set so as to effect system performance in the worst possible manner. The system is designed to operate satisfactorily under these hostile conditions. However, as pointed out by Hellerman and Racite (17), these methods have two serious shortcomings. First, the event that all critical component characteristics will simultaneously attain limiting values in actual operation is extremely small so that the components have unnecessarily tight tolerances; second, they provide no measure of relative reliability so that comparative costs for different situations cannot be weighed against increases in operational worth.

A non-redundant system is one in which the catastrophic failure of any component causes system failure; i.e., every component is essential to proper operation. The effects of catastrophic failure in a non-redundant system can be reduced in either of two ways: (1) by using more reliable components, (2) by using redundancy. With systems containing a large number of components, the first is often ineffective. However, Davis (18) cites a case when redundancy is impractical because of space, weight, and cost considerations. On the other hand, Luebbert (19) contends that redundancy is the only realistic approach to the failure problem.

Redundant components are components which are extraneous if all other components are working properly, but which can replace catastrophically failed components without interrupting the operation of a system. Three types of redundancy are distinguished, called active, standby, and majority redundancy.

If a redundant component is used continuously from the time the system is energized until failure, the system contains active redundancy. If a failed component is automatically replaced by a new component, using a detection and switching apparatus, the system contains standby redundancy. Flehinger (20) has presented a very thorough analysis of standby redundant systems.

These two types of redundancy are not both applicable to all situations. For instance, the total resistance of a set of paralleled resistors changes when the separate resistors fail. However, two electrical generators can form an active redundant

system, if they are paralleled and both have the proper loading capacity. Alternately, they can be used in standby, if a detection and switching mechanism is available. Both active and standby redundancy are used with continuous, repairable or non-repairable, systems.

Majority, or voting, redundancy was introduced by Von Neumann (21) and extended by Depian and Grisamore (22). The primary application of majority redundancy is to gate circuits in digital computers, which are intermittent systems. A single gate is replaced by a set of identical gates whose binary outputs are polled by a majority device. The output of this device is designed to agree with the majority of inputs.

One of the few design procedures for constructing a redundant system was presented by Moore and Shannon (23, 24) in their classical study of "crummy" relays. Active redundancy is used, but the systems are intermittent.

Only active redundant systems are treated in this thesis, although the failure analysis can easily be extended to standby redundant systems.

CHAPTER II

MATHEMATICAL MODELS OF FAILURE ANALYSES

Every analytical paper treating the influence of component failure on system performance introduces, at least implicitly, a reference frame, or set of "ground rules" which acts as a model and upon which the paper's computations are based. In the great majority of such models, failure types, distributions, or component interconnections are specified, and are used in the computation of numerical figures of merit. A comprehensive mathematical model which contains these detailed models as special cases is outlined in the first section.

This broadly based mathematical model is termed the General Model. It embraces many types of failure analyses yet permits direct computation of figures of merit. With this model, three features are incorporated into figures of merit which are not apparent in any model available in the literature. First, both the drift and catastrophic failure tendencies of a system are combined; second, the drift characteristics of components in redundant systems are included; third, the changes in the failure-causing environment of a component caused by the failure of other components is taken into account.

The salient aspects of three important models presented in the literature and their relationships to the General Model are

presented in the second section.

Foundations of the General Model

The General Model contains a set of definitions which provides the framework for statistically describing the components of a system, for relating these descriptions to system behavior, and for evaluating corresponding figures of merit. Definitions of the mathematical terms used to construct the General Model are given in the first part of this section. In the second, some measures of operational effectiveness are defined. The procedural technique for using the General Model to carry out a failure analysis is presented in the third part of this section.

Definitions

Part parameter

For the purpose of analysis, a system can be broken into units, usually called subsystems and components. The most basic units which can be defined are called irreducible multiterminal components¹, each of which is completely characterized mathematically by a terminal representation. Any coefficient in the corresponding terminal equation is called a part parameter if it satisfies the following two conditions.

- i) It is a random variable with a non-trivial distribution.
- ii) It is not a function of other random variables.

Any through or across variable which satisfies these two conditions is also termed a part parameter.

The set of part parameters defined for a system must exhibit all the inherent statistical characteristics of the system's

¹This notation has been introduced by Koenig and Blackwell (25, p. 109).

building blocks.

Each part parameter is represented by a stochastic process¹, written $(B_i(t), t \in T_i)$ where T_i is a linear point set representing usage times. The values of $B_i(t)$, for any t , are separated into two mutually exclusive sets, labelled $C_i(t)$ and $D_i(t)$.

The set $C_i(t)$ is assumed to contain at most two points, c_{i1} and c_{i2} which are extreme values of $B_i(t)$, both of which connote complete breakdown or catastrophic failure. A part parameter is said to have failed at time t if $B_i(t) \in C_i(t)$. Since many part parameters may be associated with a component, the relationship between failure in the real sense and a part parameter assuming an extreme value is often nebulous. Thus, a physical interpretation of part parameter failure will not be made. Failure is permanent so that if $B_i(t) = c_{ij}$, then $B_i(\tau) = c_{ij}$ for all $\tau > t$.

The set $D_i(t)$ contains all possible drift values. If $B_i(t) \in D_i(t)$, the part parameter may or may not be operating properly, depending on tolerance limits.

The set of all part parameters is represented by the vector $\bar{B}(t) = (B_1(t), \dots, B_k(t))$.

Parameter state variable

A convenient format for expressing the distribution of $B_i(t)$ is afforded by an auxiliary random variable $X_i(t) = X(B_i(t))$, called the parameter state variable, or the state variable for the part parameter $B_i(t)$. The range of this random variable is

¹The definition of a stochastic process and the notation used are given in Appendix A.

defined in Eq. (1).

$$\begin{aligned}
 X_i(t) &= 0 \text{ if } B_i(t) \in D_i(t) \\
 (1) \quad &= 1 \text{ if } B_i(t) = c_{i1} \\
 &= 2 \text{ if } B_i(t) = c_{i2}
 \end{aligned}$$

The numerical values assigned are, of course, entirely arbitrary.

The distribution of $B_i(t)$ can then be prescribed by the cumulative distribution function $F_{B_i}(b;t)$ in Eq. (2).

$$(2) \quad F_{B_i}(b;t) = \sum_{x=0}^2 F_{B_i/X_i}(b/x;t) p_{X_i}(x;t)$$

In Eq. (2), $F_{B_i/X_i}(b/x;t) = p(B_i(t) \leq b / X_i(t) = x)$ and $p_{X_i}(x;t)$ is the density function of $X_i(t)$.

The convenience of this notation becomes apparent when the individual terms are considered. $F_{B_i/X_i}(b/0;t)$ represents the drift characteristics of the part parameter. $F_{B_i/X_i}(b/1;t)$ and $F_{B_i/X_i}(b/2;t)$ are step functions, each with a saltus of one. The density factor expresses the relative probability of catastrophic and drift values.

System state variable

If the drift characteristics of a redundant system are to be studied at a certain time, both the part parameters which have not failed and their connections must be known. This information is supplied by an auxiliary random variable, called the system state variable, denoted by $X(t)$. Depending on the order in which the part parameters fail and the type of redundancy, different drift situations can arise. A one-to-one correspondence is established between the values of the system state variable and

all possible drift situations, or states. Specific values for the system state variable are assigned in later sections when active redundancy is considered. In the following discussion, the system state variable is assumed to have distinct values x_1, \dots, x_m .

System parameters

A system parameter is any function of part parameters used to make a quantitative judgement on the excellence of performance.

Each system parameter is represented by a stochastic process, $(A_i(t), t \in T_i)$. The set of all system parameters is denoted by the vector $\bar{A}(t) = (A_1(t), \dots, A_n(t))$.

The relationship between the i^{th} system parameter, the set of part parameters $\bar{B}(t)$, and the system state variable $X(t)$ is given in functional notation by Eq. (3).

$$(3) \quad \begin{aligned} A_i(t) &= h_i(\bar{B}(t), X(t); t) \\ &= h_{ij}(\bar{B}(t)) \text{ for all } t \text{ such that } X(t) = j. \end{aligned}$$

The form of Eq. (3) indicates that each system parameter can become a function of different part parameters when the system changes its state; i.e., when part parameters fail.

The (continuous) density function for the joint distribution of the system parameters is expressed by Eq. (4).

$$(4) \quad f_{\bar{A}}(\bar{a}; t) = \sum_{j=1}^m f_{\bar{A}/X}(\bar{a}/x_j; t) p_X(x_j; t)$$

The cumulative distribution function corresponding to Eq. (4) is denoted by $F_{\bar{A}}(\bar{a}; t)$.

If a large system is divided into a set of components, the

contribution of each component to over-all operation is expressed by a set of component parameters. They play the same role for the components as the system parameters play for the system. The system parameters for the large system are then expressed in terms of these component parameters. If some of the components are labelled as subsystems, their parameters are called subsystem parameters.

Value functions

All system characteristics to be evaluated quantitatively are contained in the system parameters. The means for computing relative numerical values for these characteristics are provided by value functions. These functions are the figures of merit by which the system's operational effectiveness is judged, using tolerance sets as criteria.

A tolerance set, denoted by $S_i(t)$ for the i^{th} system parameter, is that set of values of system parameter $A_i(t)$ such that the statement " $A_i(t) \in S_i(t)$ at time t " means that that portion of system performance judged by $A_i(t)$ is completely satisfactory at time t . The tolerance sets are defined in terms of design objectives and limits imposed by specifications.

The possible value functions are many and varied. This thesis is concerned with two of these, called reliability and utility which are defined below.

Reliability

The standard definition of the reliability of a system has been given by many authors (e.g. Bazovsky (26, p. 11)). "Reliability is the probability of a device performing its purpose

adequately for the period of time intended under the operating conditions encountered."

Adequate performance is taken to mean performance within the tolerance sets. Then, the reliability, $R(t_0, t_1)$, of a system for the time interval (t_0, t_1) , $t_0 < t_1$, is defined by Eq. (5).

$$(5) \quad R(t_0, t_1) = P(A_i(\mathcal{T}) \in S_i(\mathcal{T}), \text{ all } t_0 \leq \mathcal{T} \leq t_1 \text{ and all } i)$$

Utility

The drift of a system parameter just outside its tolerance set might be much less serious than an extreme value. This effect is included by defining a utility function, written $u_i(A_i; t)$, for each system parameter. The range of each parameter utility function is arbitrarily taken as $[0, 1]$ with the maximum value shown below.

$$u_i(A_i; t) = 1 \text{ if } A_i(t) \in S_i(t)$$

This assigns maximum utility when the system is operating as well as can be expected, as far as the i^{th} system parameter is concerned. When $A_i(t)$ assumes an extreme value, u_i might be taken as zero. The values of u_i for system parameter values between these limits might be taken as a non-increasing function of $A_i(t)$.

Some system parameters might reflect more essential properties than others. This is taken into account with a weighted sum of parameter utility functions, called the system utility function $U(\bar{A}; t)$ defined in Eq. (6).

$$(6) \quad U(\bar{A}; t) = \sum_{i=1}^n w_i(t) u_i(A_i; t)$$

The following conditions are assumed to be satisfied in Eq. (6).

$$\sum_{i=1}^n w_i(t) = 1 \text{ and } w_i(t) \geq 0, \text{ any } t \text{ and } i = 1, \dots, n.$$

The weights are made functions of time so that the relative importance of the various parameter utility functions can be changed in time. The system utility function has range $[0,1]$. Measures of the distributions of the parameter utility function and of the system utility function, such as the mean and variance, serve as numerical values of utility.

Failure analysis using the General Model

A simple illustration of a multi-loop control system will show how the operation of a system is viewed in terms of the General Model. Both the forward and feedback loops might contain redundant components.

A set of system parameters such as response time and gain margin, are defined to measure the operational effectiveness of the system. These parameters are functions of component characteristics such as plant constants, amplifier gains, and gear friction. All such characteristics which are assumed to be random variables are designated as part parameters.

The system is turned on at time zero. The part parameters drift from their initial distribution as the system operates, causing changes in the distributions of the system parameters. When a part parameter fails, the system changes states. The effect of this failure could be minor, if proper redundancy is

available, or it could remove a feedback loop. This, in turn, could simply degrade system performance or could cause instability and system failure, depending on the type of failure and number of feedback loops. The amount by which the system's operational worth is decreased depends on the tolerance sets assigned to the system parameters and on the choice of value functions.

By a failure analysis is meant the determination of value functions whose numerical values depend on usage time and on the distributions of part parameters. The five steps below outline such a failure analysis on the basis of the General Model.

1. Select the part and system parameters. Because of analytical complexity, all component characteristics may not be designated as part parameters. The choice depends on the inherent variability in the component characteristics and on the sensitivity of system performance to these variations. The set of system parameters must express all facets of system operation which are to be examined, as dictated by design objectives.

2. Determine the distribution of the system state variable. Both the range and distribution of the system state variable must be found. This step is the main topic of Chapter III.

3. Perform a drift analysis for each state. The joint distribution of the system parameters for each state of the system is found. Since the system changes states when a part parameter fails, the forms of the system parameters change, as shown in Eq. (3). Chapter IV covers the execution of this step.

4. Combine the drift and catastrophic failure effects. The results of Steps 2 and 3 are combined, using Eq. (4), to find the joint distribution of the system parameters.

5. Determine the value functions. The joint distribution derived in Step 4 is used to form the reliability and utility functions, both of which depend on usage time and part parameter distributions.

In summary, the cornerstone of a failure analysis using the General Model is the concept of separating the analysis into a set of problems, each of which can be solved independently. The system state variable provides the means both for this separation and for the recombination of these problems into the over-all analysis and applies the proper weight to each result. That is, drift failures in more probable situations are weighted more heavily than those in less probable situations. The catastrophic failure tendencies of the part parameters determine the distribution of the system state variable. Thus, both types of failures are packaged together so that operational effectiveness can be judged on the basis of all pertinent factors.

Although the five basic steps are used in all failure analyses, their execution can take on many forms. The entire analysis may be accomplished analytically or a computer-simulated sampling plan can be set up. The computational details depend on the interaction between part parameters, the types of failure distributions, and the forms of the system parameters. Also, the analysis might be simplified into finding certain measures of the distributions of system parameters.

It should be emphasized that the General Model is not limited to reliability and utility value functions, but provides the foundation for calculating any analytical value function.

Existing models

In this section, the salient features of three mathematical models from the literature having wide applications are presented, along with their relation to the General Model. However, before these three models are discussed, some other work on model building deserves mention.

Two papers by Rosenblatt (27, 28) present, somewhat qualitatively, some features which are desirable in a model. A series of simple models are given along with methods for expressing physical events in probability language.

An early approach to an over-all failure study was presented by Benner and Meredith (29). The assumptions made include normally distributed performance parameters, series systems, and independently failing components. This frequently-referenced paper presented ideas which were quite novel at the time of its publication.

A theoretical exposition of the factors causing component failure has been given by Birnbaum (30). The physical condition of a structure is considered as a stochastic process whose properties depend on environment, usage, and observed samples. However, the model is not directly applicable to the calculation of value functions.

Many facets of a failure analysis are discussed in a paper by Drenick (16). Most of the techniques presented are applications of renewal theory. However, cost minimization procedures are given for the marginal checking technique. Several figures of merit are used. A rather quantitative method for handling drift

is considered. This paper is more a collection of techniques than a comprehensive working model.

Renewal models

A brief outline of the aims of renewal theory applicable to failure analyses is presented below. The following assumptions are made.

1. All system components must be operating properly for the system to be operating properly.
2. Removal of a system component is caused either by failure during operation or by a predetermined maintenance policy.
3. Upon removal, a component is immediately replaced by a new and statistically identical one.

According to Drenick (31), the system may be viewed as a number of sockets into each of which a component is inserted. At each socket, a renewal process develops, i.e., an unending sequence of removals and replacements. Corresponding to each socket, there is a parent population from which a component is drawn each time a replacement is required. The over-all failure pattern for the system is then a superposition of many such renewal processes. The derivation of this failure pattern's properties is the main objective of a renewal study.

Two functions, called the renewal rate and the expected number of replacements, relate the renewal process at each socket and the system removal pattern. These functions are found from the solution to differential equations and depend on the parameters of the component failure distributions.

The above assumptions imply that a component failure can be

detected, located, and corrected in a negligible period of time; i.e., the system is always operating. Implied here are the assumptions that system failure can be traced to one particular component and that replacement of that component restores the system to proper working order. Failure caused by component drift is not allowed, so that this theory is directly applicable only to catastrophic failures.

The case when all components have the same lifetime distributions has been studied in some detail by Flehinger and Lewis (32). One interpretation of this is that all system components have the same parent population. Detailed solutions to the resulting renewal equation, using mean-time-to-failure and survival probability as figures of merit, are also presented for various hazard functions.

The system failure pattern when the renewal processes for all the sockets are not identical has been studied by Drenick (31) in an approach motivated by the central limit theorem. An asymptotic expression for survival probability of the form $\exp(-t)$ is derived under certain conditions. An interesting feature of this approach is that the asymptotic expression becomes more accurate as the number of sockets increases. A restrictive assumption is that of independence of the renewal processes for the various sockets.

It is stretching a point to say that the renewal model is a special case of the General Model. If each part parameter is allowed only two values, and if the three assumptions are applied, a study of the General Model implies a study of the renewal

model. However, both the mathematical techniques used and the aims of the investigations are different. Although some features of the renewal model at the component level will be incorporated into the General Model, the viewpoint taken is that the two models cover different areas of study and should not be combined.

Markov chain models

If drift is neglected and if all component failure rates are constant, the theory of stationary Markov chains¹ can be applied directly to a failure study. A most interesting analytical model using this approach has been presented by Barlow and Hunter in two papers (33, 34). The first paper (33) states the mathematical and probabilistic basis of the model. The second (34) is a special case which covers a type of optimal redundancy for series-parallel systems along with maintenance and checking policies. The optimal redundancy study has also been published in a separate paper (35). The main features of the Barlow-Hunter model are discussed below.

The essential description of a system is given in terms of a stochastic process which describes the system's state. The number of states is determined by the number of components, their connections, and the number of states each component can assume. In all examples presented, each system component can, at any time, be in one of two or three states. One corresponds to "in use"; the others, to "in-repair", where both open and short circuit types of failures are allowed. Time-to-failure

¹An outline of the properties of continuous parameter Markov chains is presented in Appendix B.

and time-to-repair are both assigned exponential distributions so that the well-developed theory of stationary Markov chains can be used. Independence of component failures need not be assumed. this technique is particularly applicable to repairable systems containing active redundancy in which drift is neglected.

Two figures of merit are used, called reliability and efficiency. Both are defined in terms of a gain function $g X(t)$ whose domain is the state space of the stochastic process and whose range is contained on the real line. Speaking qualitatively, the value of the gain function should be relatively large for favorable states of the system and relatively small for unfavorable states.

The following definition of reliability is used by Barlow and Hunter.

$$R(t) = E[g(X(t))] = \int g X(t,w) dP(w)$$

A particularly useful gain function is expressed in terms of a class A of favorable states of the system.

$$\begin{aligned} G X(t) &= 1 \text{ if } X(t) \in A \\ &= 0 \text{ otherwise} \end{aligned}$$

With the gain function G, the R(t) function becomes the probability that the system is in a favorable state at time t.

$$R(t) = \int_A dP(w)$$

The second figure of merit, called the efficiency and denoted by Eff, is defined below.

$$\text{Eff} = \iint g X(t,w) dP(w) dF(t)$$

In this expression, $F(t)$ is the distribution function for a time-based distribution which could reflect the presence of an environmental threat in time. If this distribution is continuous and uniform on $[0, T]$, Eff is the time average of $R(t)$ over that interval.

Two observations should be made concerning the definitions of reliability used by Barlow and Hunter.

1. Since reliability is here defined as the expected value of a random variable, it is not, in general, a probability. No restrictions have been placed on the range of $g[X(t)]$, so its value is not limited to the unit interval, or even to positive values. If the gain function given above is always used, this objection is eliminated.

2. Even with the special gain function, the $R(t)$ function is not reliability in the usually-accepted sense, since it is not a function of a time interval. The objection here is not to the usefulness of the figure of merit, but to its name.

The Barlow-Hunter Model is a special case of the General Model. Assuming that each component can be characterized by a single part parameter, the two models are related by properly interpreting the two notations. If drift is neglected, the parameter state variables completely describes the failure tendencies of the part parameters. Thus, the parameter state variable and the state variable for each component used by Barlow and



Hunter are synonymous. The three values assigned to the parameter state variable in Eq. (1) correspond to the allowable states for each component in the Barlow-Hunter Model. The states of the system are the same in both models. The gain function and the utility function serve the same end. The $R(t)$ function and the efficiency are then measures of utility.

Another approach using Markov chains has been presented by Rohn (36). Standby, rather than active, redundancy is the main concern of this paper. The figure of merit used is called "fractional up-time" which is the fraction of time the system is operable during a time interval. Drift failure is neglected and only one mode of catastrophic failure is allowed. Variations in the operational worth of the various standby components are included. Both time-to-failure and time-to-repair are assigned exponential distributions. Ashar (37) has also studied this problem using Markov chains, but from a more general viewpoint.

Roberson's model

One of the only models in which both catastrophic and drift failures are included in a unified analysis has been presented by Roberson (38). The essential features of this model are discussed below.

System performance is measured by a vector \bar{x} of scalar quantities representing variables such as voltage, pressure and counter readings. Ideal, or perfect, system performance is defined by a value of \bar{x} , denoted by \bar{x}^* . System error, e , is defined as the norm of $(\bar{x} - \bar{x}^*)$ and measures the system's operational worth in terms of a value function which is a

non-increasing function of system error.

The value function to which most of Roberson's paper is devoted is called "p-error", and is defined as the solution to the following equation.

$$P(e \leq e_p) = p \quad (0 \leq p \leq 1)$$

The distribution of the system error is defined in terms of the zero-failure probability R , the inherent accuracy, $A(e')$, and the failed accuracy, $F(e')$.

$$R = P(\text{no component failure})$$

$$A(e') = P(e \leq e', \text{ if no component failure})$$

$$F(e') = P(e \leq e', \text{ if one or more component failure})$$

The zero-failure probability R equates system failure to failure of any component. The parameters of A are called accuracy parameters and describe the drift properties of the system when all components are operating properly. Different degrees of operational worth are inserted for varying degrees of system failure by means of the parameters of F , called failure severity parameters.

The p-error satisfies the following equation.

$$p = R A(e_p) + (1-R) F(e_p)$$

The value of p establishes a performance level. If a value of p is specified, the value of e_p can be found from the above equation. Thus, the relationship between probable system error and the system performance level can be derived. Roberson discusses in some detail the existence of and bounds on solutions for e_p .

The main limitation on this model is that it is applicable

only to series systems. This is implied by the zero-failure probability R which includes all catastrophic tendencies. This measure also allows for only one mode of catastrophic failure.

Roberson's model is a special case of the General Model. The vector \bar{x} corresponds to the part parameters; system error e serves as the system parameter. If norm $(\bar{x} - \bar{x}^*)$ is identically zero when $\bar{x} = \bar{x}^*$, the tolerance set on e is the single point zero. Using a single set of values as ideal performance provides a direct relationship between the tolerance set on the system parameter and limits on the part parameters. However, it seems more desirable to speak of a range of values as satisfactory performance. The distribution constants for the distributions of the part and system parameters serve the same purpose as the accuracy and failed-severity parameters used by Roberson. The p-error and the utility function measure the same characteristics of system behavior.

CHAPTER III

CATASTROPHIC FAILURES

Methods for determining the range and distribution of system state variables for systems containing active redundancy are presented in this chapter.

As a by-product of this study, techniques for finding reliability expressions for active redundant systems in which drift is neglected are derived. Two results are presented which have not been given in the literature. First, an algorithm for deriving reliability expressions for non-loaded systems is given which is independent of failure distributions and component connections. When properly interpreted, this algorithm can be used when the part parameters have one or two failure modes. Second, a method for determining the reliability of any loaded system composed of exponentially failing part parameters is described.

In the first section, systems in which each part parameter has one catastrophic failure mode are discussed; in the second, two failure modes are allowed.

One catastrophic failure mode

As was discussed previously, the values of the part parameter $B_i(t)$ are segregated into two sets, $D_i(t)$ and $C_i(t)$, which are interpreted as drift and catastrophic values,

respectively. In this section, $C_i(t)$ is assumed to be a one-point set c_i . The value c_i is either of the two extreme values allowed for $B_i(t)$. For notational convenience, these two extreme values are referred to as the open and short modes of failure.

Range of the system state variable

Before the states of a system, or values of the system state variable, can be enumerated, the concept of catastrophic system failure must be defined. A system has failed (catastrophically) when it is considered to be operationally useless. This state could correspond to a system parameter value far outside its tolerance set or to an extreme value of a value function. When in this state, the system is shut down; no further part parameter failures can occur. A physical interpretation of system failure can be made only on a particular system.

The states of a system are designated in terms of the parameter state variables. A numbering system for accomplishing this is presented below.

The base r representation of any (base 10) non-negative integer N_{10} can be written as $M_k M_{k-1} \dots M_1$ if Eq. (1) is satisfied.

$$(1) \quad N_{10} = \sum_{i=1}^k M_i r^{i-1}, \quad 0 \leq M_i \leq r-1$$

In particular, if $r=2$ and if Eq. (2) is satisfied, then $M_k M_{k-1} \dots M_1$ is the binary representation of N_{10} .

$$(2) \quad N_{10} = \sum_{i=1}^k M_i 2^{i-1}, \quad M_i = 0 \text{ or } 1$$

The range of a parameter state variable $X_i(t)$ was defined in Eq. (II.1). In this section, $c_{i1} = c_i$ and c_{i2} is a set of measure zero. Assuming that the system is turned on at $t = 0$, at which time no part parameter has failed, state numbers are assigned by the following two rules.

1. If, at time t , the system has not failed, let $M_i = X_i(t)$. Let the number N_{10} obtained from Eq. (2) be the state number at t .

2. Consider all numbers formed as in Rule 1 for which the system is failed. Let the minimum of these be the state number corresponding to system failure.

The binary representation of each state number will be written as a k -digit number, assuming the system has k part parameters. If a system is in a non-failed state at time t , such a representation exhibits the values of the parameter state variables at t ; the left-most binary digit, or bit, is the value of $X_k(t)$, the second from the left, $X_{k-1}(t)$, and so forth so that the right-most bit is the value of $X_1(t)$.

Two questions must be answered before the two rules can be used to enumerate the states of a particular system: (1) which sets of part parameters can fail before the system fails; (2) which sets of part parameters can be in the failed state at the same time? This information can be presented in either of two forms--by a state table or by a reliability diagram.

The state table exhibits all sets of failed part parameters and the corresponding state numbers in tabular form. A state table has $k+1$ column headings, one for each parameter state variable and one for the system state variable. Although the headings can be listed in arbitrary order, it is convenient to write them in the order $X_k(t) X_{k-1}(t) \dots X_1(t) X(t)$. The numbers 0 to 2^{k-1} are listed in binary form under the first k headings and the state numbers are determined by the two rules. If the system is operational for the values in a particular row, the state number of the system is the base 10 equivalent of that row's binary number. All rows which represent impossible states are omitted.

An example of a state table and corresponding reliability diagram is given in Fig. 1 in which all failures are assumed to be of the open variety.

The state table can be generated in the following manner. Consider any set of values for the parameter state variables. If $X_i(t) = 0$, let $B_i(t)$ be its nominal value; if $X_i(t) = 1$, let $B_i(t) = c_i$. For each possible set of part parameter values, determine if the system has failed. Then, apply the two rules for finding state numbers. This is repeated until the state table is completed. A set of values can be ignored if the corresponding event has measure zero; i.e., the failure of one part parameter might effectively remove other part parameters from the system by shutting down some subsystem. This process can be programmed on a digital computer.

A reliability diagram contains the same information as a

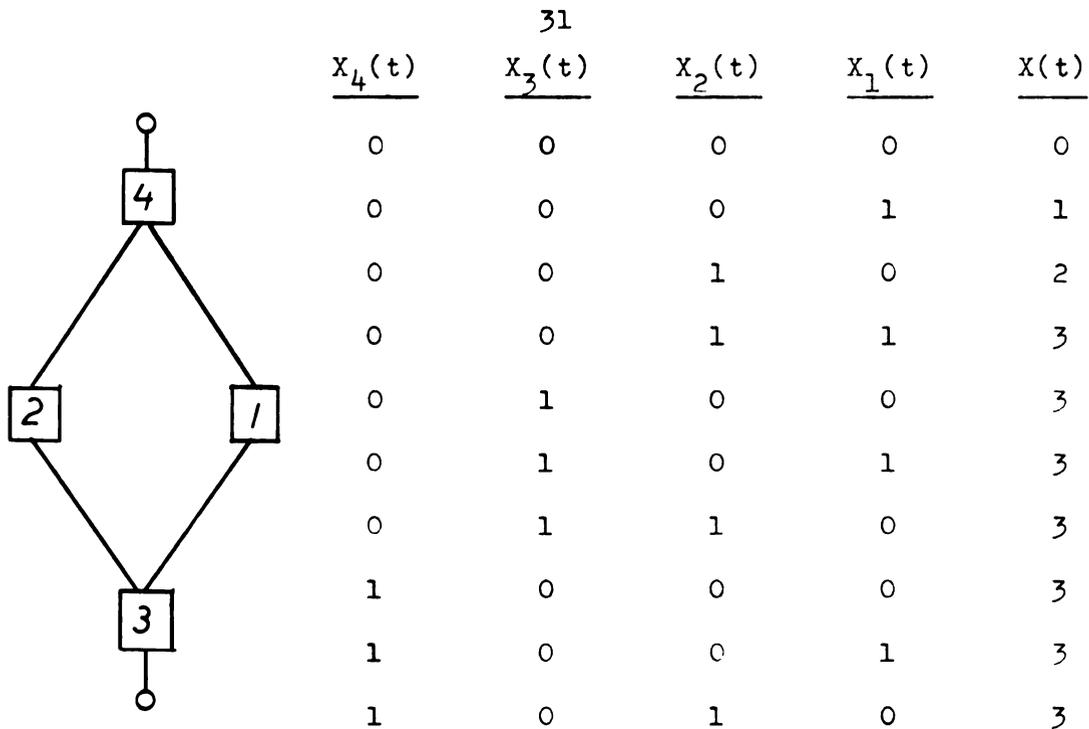


Fig. 1.--Example of a state table and corresponding reliability diagram.

state table, except that it does not number the states explicitly. It merely exhibits schematically the relation between part parameter and system failures. A step-by-step procedure for constructing a reliability diagram will not be given. Rather, a set of rules by which an existing diagram can be interpreted is presented, followed by some examples.

As used in this thesis, a reliability diagram is a collection of blocks, one for each part parameter, and lines drawn between two vertices. The lines connect the blocks so that each block is contained in at least one path between the vertices. The block for each part parameter is labelled with that parameter's subscript. The reliability diagram is interpreted by means of the following rules.

1. If $B_i(t) = c_1$, i.e., $X_i(t) = 1$, the i^{th} block is

either erased from the diagram (open failure) or replaced by a solid line (short failure) depending on the single mode of catastrophic failure assigned to $B_i(t)$.

2. The system is in the failed state either if a solid line exists between the two vertices, or if there is no path consisting of blocks and lines between them. Otherwise, the system is operational. Some simple examples are shown in Fig. 2.

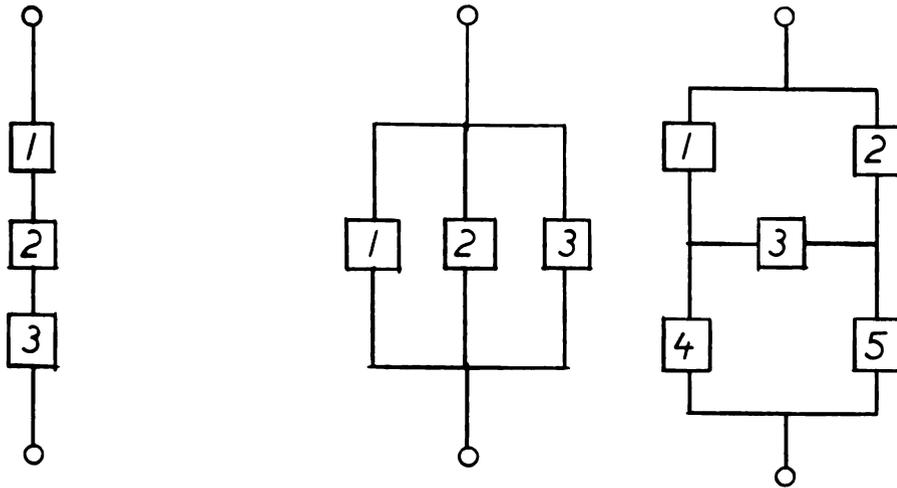


Fig. 2.--Examples of reliability diagrams

In Fig. 2A, system operation ceases if any of the three part parameters fails open, or if all fail short. The situation is reversed in Fig. 2B; all must fail open or any one can fail short for system failure. Fig. 2C is a type of bridge diagram. If all part parameters can fail in the open mode, system operation ceases if all part parameters in any of the following sets are failed: (B_1, B_2) , (B_4, B_5) , (B_1, B_3, B_5) , (B_2, B_3, B_4) , (B_1, B_2, B_3) , (B_3, B_4, B_5) . It should be noted that, for example, the event $(B_i(t) = c_i, \text{ all } i=1,2,3,4)$ has measure zero.

If all part parameters in Fig. 2C can fail only in the

short mode, the catastrophic failures of all the part parameters in any of the following sets would cause system failure:

(B_1, B_4) , (B_2, B_5) , (B_2, B_3, B_4) , (B_1, B_3, B_5) , (B_2, B_3, B_5) , (B_1, B_3, B_4) .

The diagrams in Fig. 2 have one feature in common--only two lines are connected to each block. Such diagrams will be called "two-line" diagrams. However, all active redundancy cannot be represented by this type of diagram, as shown in Fig. 3.

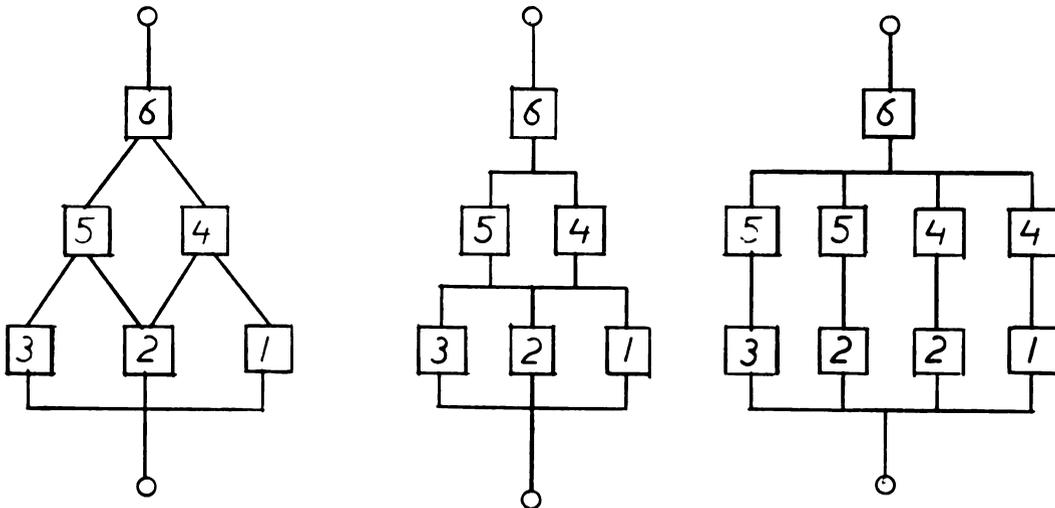


Fig. 3.--Reliability diagrams

At first glance, Fig. 3A and Fig. 3B appear to represent the same system. However, blocks 4 and 5 are in parallel in Fig. 3B, but are not in Fig. 3A. Also, if all part parameters except B_1 , B_5 , and B_6 , failed open, the system represented by Fig. 3B is operable, while that represented by Fig. 3A has failed. It might seem that the diagram in Fig. 3A can be transformed into an equivalent two-line diagram, such as that in Fig. 3C, since they have the same blocks in the paths between the vertices. However, the two diagrams do not produce the same

reliability number.

The type of diagram drawn in Fig. 3A is, then, more general than a two-line diagram. It sometimes has a closer relationship to block diagrams for physical systems and should be easier to construct. However, in general, there is no direct correspondence between the topology of a reliability diagram and those of signal flow graphs, linear graphs, or schematic diagrams.

The primary difference between the variety of reliability diagrams proposed in the literature (39, 40) and those used in this thesis is the interpretation given each block. In this thesis, the blocks represent part parameters, while in the usual diagrams, they represent components.

Distribution of the system state variable

The concept of a lifetime random variable is presented first, followed by a discussion of loading effects. The process of state changing is then discussed. This subsection is concluded with a review of the analytical techniques available for deriving the density function for the system state variable.

Lifetime random variables

The catastrophic failure tendencies of a part parameter are often exhibited by means of an auxiliary random variable, called the time-to-failure, or mortality, or lifetime random variable, denoted by L . The range of each lifetime variable is the usage time of the system. The event $[L_i = t]$ means that the i^{th} part parameter fails at time t . The range of the parameter state variable for part parameter $B_i(t)$ is expressed in terms of the lifetime variable in Eq. (3)

$$(3) \quad X_i(t) = 0 \text{ if } L_i > t \quad (\text{no failure up to time } t) \\ = 1 \text{ if } L_i \leq t \quad (\text{failure at or before } t)$$

The alternate expression for the range of a parameter state variable in Eq. (3) is introduced to show how the well-known concept of time-to-failure ties in with the General Model. The lifetime variable is especially useful when the reliabilities of systems in which drift is neglected are studied. The relationship between the distributions of part parameters, their state variables, and lifetime variables is given in Eq. (4).

$$(4a) \quad P(L_i \leq t) = P(B_i(\mathcal{T}) = c_i, \text{ some } \mathcal{T} \leq t) = P(X_i(\mathcal{T}) = 1, \\ \text{some } \mathcal{T} \leq t) \\ = P(X_i(t) = 1) = p_{X_i}(1;t)$$

$$(4b) \quad P(L_i > t) = P(B_i(\mathcal{T}) \neq c_i, \text{ all } \mathcal{T} \leq t) = P(X_i(\mathcal{T}) = 0, \\ \text{all } \mathcal{T} \leq t) \\ = P(X_i(t) = 0) = p_{X_i}(0;t)$$

Loading effects

The technique used to find the distribution of a system state variable depends to a great extent on whether the lifetime distributions change after the system's state changes. When a part parameter fails, the stresses on the remaining part parameters might increase, thus increasing the chance of their failing. The amount by which a part parameter is weakened could depend both on which states the system has passed through and on how long the system has remained in each of these states. Such effects are referred to as loading effects.

A system is classified either as non-loaded or as loaded. The set of all states for which the system is operational is denoted by Z_0 . If, for all $x^* \in Z_0$, the joint density function $p_{\bar{X}}(\bar{x}; t)$ of the parameter state variables is given by Eq. (5), the system is said to be non-loaded. The binary representation of x^* is written as $x_k^* \dots x_1^*$.

$$(5) \quad p_{\bar{X}}(\bar{x}^*; t) = \prod_{i=1}^k p_{X_i}(x_i^*; t)$$

It should be noted that the parameter state variables in a non-loaded system are not statistically independent, since no part parameter can fail after the system has failed.

The operation of a non-loaded system can be viewed in terms of the following illustration. Consider each part parameter as being operated at a separate test station, rather than in the system. However, let the failure-causing environment of each part parameter be the same as would be experienced in the system. As the part parameters fail, consult either the reliability diagram or the state table to determine when the system fails. Shut off all test stations at the instant of system failure.

A non-loaded system can be considered as one in which failures are caused mainly by the system's external environment such as vibration, humidity, and temperature. In a loaded system, failure is also caused by the task the system is performing, e.g., by the currents, forces, or pressures associated with the various part parameters.

The failure pattern in a loaded system is obviously more complicated than in a non-loaded system. When loading is a factor, the illustration cited above is changed. The environment at each test station varies with time and depends on which part parameters have failed, the order in which they failed, and how long each part parameter had operated.

The state-changing process

It will be assumed that each lifetime variable is of the continuous type. Then, the event in which two or more part parameters fail simultaneously has measure zero. Thus, a system can go from state i to state j without passing through intermediate states if and only if the binary representations of i and j differ by one bit, assuming that neither state i nor state j is the failed state. Since repair is not allowed, the bit must be a 0 in i and a 1 in j , which means that the part parameter corresponding to that bit has failed.

If state j is the failed state, its binary representation may differ from that of i in any number of bits because a common state number was assigned to all situations implying system failure. For example, in Fig. 1, state 1 goes into state 3 if any one of its 0 bits changes to a 1.

When finding a system state variable's distribution, it is often convenient to record all possible inter-state transitions on what will be called a transition chart. A one-to-one correspondence is established between the possible states of the system and a set of nodes. If a one-step transition from state i to state j is possible, a directed line is drawn from node i to

node j. All possible transitions are so recorded. An example of a state table and transition chart is given in Fig. 4 in which all failures are in the open mode.

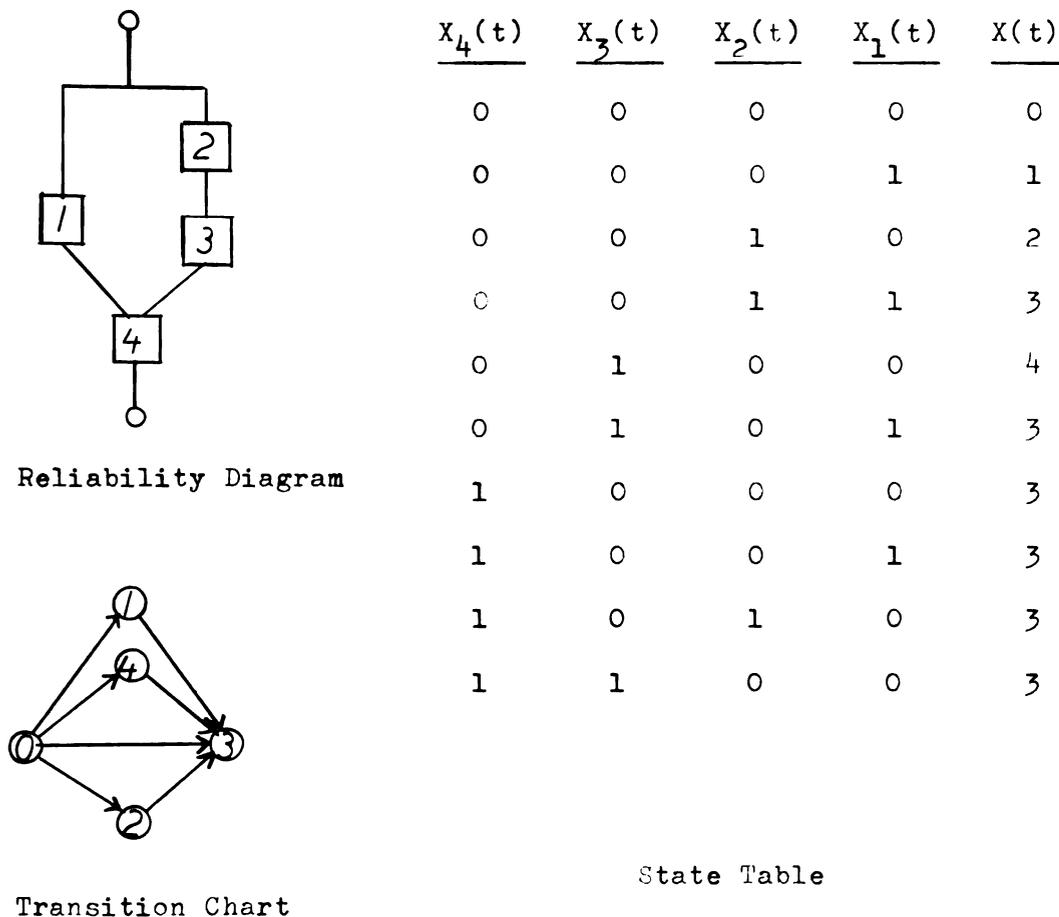


Fig. 4.--Example of a state table, transition chart and reliability diagram.

Distribution of the system state variable for non-loaded systems

A technique for deriving the density function of the system state variable for non-loaded systems is given below. The set of all state numbers for which the system is not failed is denoted by Z_0 ; the failed state, by Z_1 . The density function of the lifetime random variable for the i^{th} part parameter is

denoted by f_i .

If $x^* \in Z_0$, where $x_k^* x_{k-1}^* \dots x_1^*$ is its binary representation, the value of the system state variable's density function, p_X , is given at x^* by Eq. (6), using Eq. (5).

$$(6) \quad p_X(x^*; t) = p_{\bar{X}}(\bar{x}^*; t) = \prod_{i=1}^k p_{X_i}(x_i^*; t)$$

The individual factors in Eq. (6) are related to lifetime distributions by Eq. (7) on the basis of Eq. (4).

$$(7a) \quad p_{X_i}(1; t) = P(L_i \leq t) = \int_0^t f_i(z) dz$$

$$(7b) \quad p_{X_i}(0; t) = P(L_i > t) = \int_t^{\infty} f_i(z) dz$$

The value of the density function at the failed state Z_1 is given in Eq. (8).

$$(8) \quad p_X(Z_1; t) = 1 - \sum_{x^* \in Z_0} \prod_{i=1}^k p_{X_i}(x_i^*; t)$$

Equations (7) and (8) prescribe the distribution of the system state variable for any non-loaded system, regardless of the form of the lifetime distributions or of component connections.

As an example, the density function of the state variable for the non-loaded system whose reliability diagram and state table are shown in Fig. 5 will be computed. All part parameters are assumed to fail in the open mode. The distribution of the i^{th} lifetime random variable, $i = 1, 2, 3$, is assumed to be

exponential; the density function f_i is as shown in Eq. (9).

$$(9) \quad f_i(z) = H_i \exp(-H_i z), \text{ if } z > 0 \text{ and the system is operational} \\ = 0, \text{ if } z < 0 \text{ or if the system is failed}$$

The constant H_i is called the hazard or the failure rate.

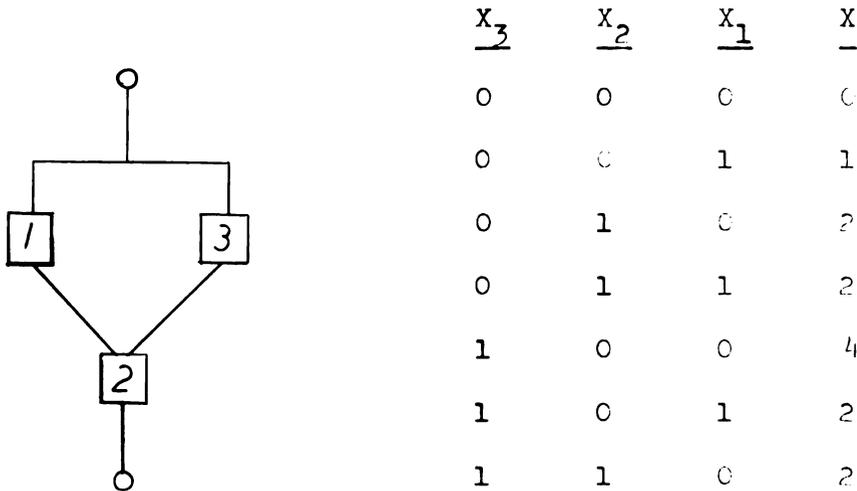


Fig. 5.--Example of a reliability diagram and state table

The following expressions are obtained from Eq. (7).

$$p_{X_i}(0;t) = \exp(-H_i t)$$

$$p_{X_i}(1;t) = 1 - \exp(-H_i t)$$

Equation (6) is used to find the density function of the system state variable at the points of $Z_0 = (0,1,4)$.

$$p_X(0;t) = p_{X_3}(0;t) p_{X_2}(0;t) p_{X_1}(0;t) = \exp[-(H_1+H_2+H_3)t]$$

$$p_X(1;t) = p_{X_3}(0;t) p_{X_2}(0;t) p_{X_1}(1;t) \\ = \exp[-(H_2+H_3)t] - \exp[-(H_1+H_2+H_3)t]$$

$$p_X(4;t) = p_{X_3}(1;t) p_{X_2}(0;t) p_{X_1}(0;t) \\ = \exp[-(H_1+H_2)t] - \exp[-(H_1+H_2+H_3)t]$$

The last value of the density function is found from Eq. (8).

$$p_X(2;t) = 1 + \exp[-(H_1+H_2+H_3)t] - \exp[-(H_2+H_3)t] \\ - \exp[-(H_1+H_2)t]$$

Distribution of the system state variable for loaded systems

Determining the distribution of the state variable of a loaded system can be a very difficult problem, since the failure times of the part parameters depend on a variety of factors. Only the case described by the following conditions will be considered.

1. At any time, the lifetime distributions for all non-failed part parameters depend only on the state the system is in at that time, and not on any state the system was in prior to that time.

2. As long as the system has not failed, the lifetime variable for each non-failed part parameter has the exponential distribution shown in Eq. (9). The hazard for each such distribution can change when the system's state changes.

These conditions reduce the problem of finding the system state variable's distribution to an application of stationary Markov chains. The notation and terminology used in Appendix B will be translated to fit the present situation.

The system state variable is described by a continuous parameter, stationary Markov process $[X(t), 0 \leq t \leq \infty]$ having a finite number of states. The states, although non-negative integers, do not necessarily include all numbers in the sequence $0, 1, 2, \dots, 2^k$.

At time zero, the system is in state 0, since no part parameters are assumed to be failed when the system is

turned on. The distribution at $t = 0$ is thus given by Eq. (10).

$$(10) \quad p_X(\mathbf{x};0) = P(X(0) = \mathbf{x}) = 1 \quad , \quad \mathbf{x} = 0 \\ = 0 \quad , \quad \mathbf{x} \neq 0$$

This implies that, at any time, the value of the density function at any state j is expressed by Eq. (11).

$$(11) \quad p_X(j;t) = p_{0j}(t)$$

As explained in Appendix B, there are essentially two methods for finding the required transition probabilities $p_{0j}(t)$. The first is to use the algorithm implied by Eq. (B.14). Equation (11) then becomes Eq. (12).

$$(12) \quad p_X(j;t) = \sum_{n=0}^{\infty} n p_{0j}(t)$$

Either of the separate induction processes given in Eq. (B.12) or Eq. (B.13) can be used to find the terms in Eq. (12). This method is particularly useful for simple topologies, such as those in Fig.'s 3, 4, and 5.

The second way to proceed is to use either the backward system of equations in Eq. (B.17) or the forward system in Eq. (B.18), together with the initial conditions of Eq. (10). The solutions to either of these sets of differential equations are the transition probabilities. This method is more applicable in large systems or where repair is allowed. Either of these systems of equations can be programmed directly for the digital computer since they are in normal form¹. Analytical

¹Digital computer solutions to differential equations in normal form have been studied by Wirth (41).

solutions can be obtained by using Laplace transform techniques and matrix algebra, or signal flow graphs.

To use either of these two methods for finding transition probabilities, the q 's in Appendix B must be expressed in terms of the lifetime distributions of the part parameters. This is discussed in Appendix C and the results are summarized below.

The lifetime density function, f_{ir} , for the part parameter $B_r(t)$ when the system is in state i is given by Eq. (C.1).

$$(C.1) \quad f_{ir}(z) = H_{ir} \exp(-H_{ir}z) \quad \text{if } z \geq 0 \text{ and } i \in Z_0 \text{ and } X_r(t) = 0 \\ \text{if } z < 0 \text{ or } i = Z_1 \text{ or } X_r(t) = 1$$

If $X_{r_m}(t)$, $m = 1, 2, \dots, n$ are all 0 when the system is in state i , then q_i is given by Eq. (C.2).

$$(C.2) \quad q_i = \sum_{m=1}^n H_{ir_m} \quad \text{if } i \in Z_0 \\ = 0 \quad \text{if } i = Z_1$$

If j can be reached from i in one step, $i, j \in Z_0$, and if $X_{r_1}(t)$ is a 0 in the binary representation of i and a 1 in that of j , q_{ij} is given by Eq. (C.3).

$$(C.3) \quad q_{ij} = H_{ir_1}$$

If $i \in Z_0$, $j = Z_1$ and if the failure of any of the part parameters $B_{r_1}(t), \dots, B_{r_v}(t)$ cause the system to fail, q_{ij} is given by Eq. (C.4).

$$(C.4) \quad q_{ij} = \sum_{m=1}^v H_{ir_m}$$

The constants q_{ij} and q_i can thus be computed from the

failure rates of the part parameters. The constants are conveniently recorded on what is called a Markov transition chart. This is the same as the transition chart defined previously, except that self loops are drawn on each node and constants are associated with each directed line. The constant q_i is associated with the self loop on the node representing the i^{th} state. The constant q_{ij} is associated with the line going from node i to node j .

To show the change in the distribution between loaded and non-loaded systems, the example of Fig. 5 will be repeated when the system is assumed to be loaded. By the notation defined above, H_{ij} is the failure rate of the j^{th} part parameter when the system is in state i . The Markov transition chart, drawn from the state table in Fig. 5, is shown in Fig. 6.

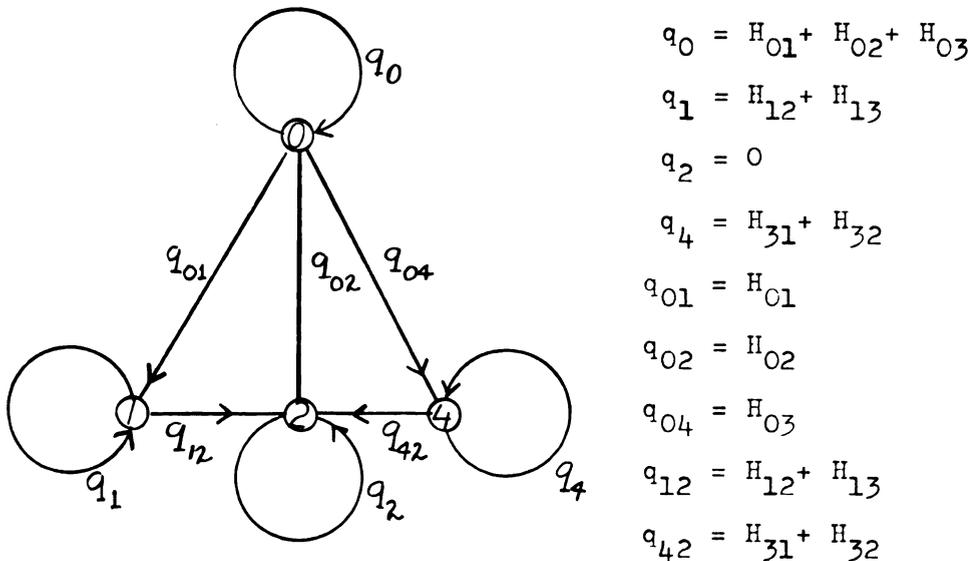


Fig. 6.--Markov transition chart corresponding to Fig. 5

The density function of the system state variable will be computed by both methods.

The first method uses Eq. (11) and Eq. (12) and the induction

process given in Eq. (B.12).

$$p_X(0;t) = p_{00}(t) = \sum_{n=0}^{\infty} n p_{00}(t)$$

$$\begin{aligned} n p_{00}(t) &= \exp(-q_0 t), & n &= 0 \\ &= 0 & , & n \neq 0 \end{aligned}$$

$$p_X(0;t) = \exp(-q_0 t)$$

Since state 1 can be reached from state 0 only by a single jump, the density function at 1 is evaluated as follows.

$$\begin{aligned} p_X(1;t) &= {}_1p_{01}(t) = \int_0^t q_{01} \exp[-q_0 s - q_1(t-s)] ds \\ &= \frac{q_{01}}{q_0 - q_1} \left[\exp(-q_1 t) - \exp(-q_0 t) \right], & q_0 &\neq q_1 \end{aligned}$$

The density function at 4 is found in the same manner.

$$p_X(4;t) = {}_1p_{04}(t) = \frac{q_{04}}{q_0 - q_4} \left[\exp(-q_4 t) - \exp(-q_0 t) \right], & q_0 &\neq q_4$$

There are three ways of going from state 0 to state 2; one way involves a single jump while the others use two jumps.

$$p_X(2;t) = {}_1p_{02}(t) + {}_2p_{02}(t)$$

The individual terms are evaluated below.

$${}_1p_{02}(t) = \int_0^t q_{02} \exp(-q_0 s) ds = \frac{q_{02}}{q_0} \left[1 - \exp(-q_0 t) \right]$$

$${}_2p_{02}(t) = \sum_{j=1,4} \int_0^t q_{0j} \exp(-q_0 s) {}_1p_{j2}(t-s) ds$$

After integrating and rearranging terms, the following is obtained.

$$\begin{aligned}
{}_2P_{02}(t) = & \frac{q_{12}q_{01}}{q_0 - q_1} \left[(1/q_1)[1 - \exp(-q_1 t)] - (1/q_0)[1 - \right. \\
& \left. \exp(-q_0 t)] \right] \\
& + \frac{q_{42}q_{04}}{q_0 - q_4} \left[(1/q_4)[1 - \exp(-q_4 t)] - (1/q_0)[1 - \right. \\
& \left. \exp(-q_0 t)] \right]
\end{aligned}$$

The expression for $p_X(2;t)$ can also be found from the following identity.

$$p_X(2;t) = 1 - \sum_{i \neq 2} p_X(i;t)$$

The second method of deriving the system state variable's density function is based on the forward system of equations given in Eq. (B.18) with initial conditions given by Eq. (10). The backward system in Eq. (B.17) could also have been used.

Since the system cannot return to a state, the forward equations need not be solved simultaneously.

$$p'_{00}(t) = -p_{00}(t) q_0, \text{ since } q_{j0} = 0 \text{ when } j \neq 0.$$

$$p_X(0;t) = p_{00}(t) = \exp(-q_0 t)$$

$$p'_{01}(t) = -p_{01}(t) q_1 + p_{00}(t) q_{01} \text{ since } q_{21} = q_{41} = 0.$$

$$p'_{04}(t) = -p_{04}(t) q_4 + p_{00}(t) q_{04} \text{ since } q_{14} = q_{24} = 0.$$

Since $q_2 = 0$, the following equation can be written.

$$p'_{02}(t) = p_{00}(t) q_{02} + p_{01}(t) q_{12} + p_{04}(t) q_{42}$$

This can be solved for $p_{02}(t)$ by direct integration,

using the previous results. The solutions to these equations for $p_{01}(t)$, $p_{02}(t)$, and $p_{04}(t)$ produce the same results as were obtained by the previous method.

It should be noted that if $H_{0i} = H_{1i} = H_{31}$, $i = 1, 2, 3$, the

system is non-loaded. In this case, the distribution derived above agrees with the previous example concerning Fig. 5.

Reliability

The General Model is now particularized to derived techniques for finding reliability expressions when drift failures are neglected. Omitting drift failures simplifies the analysis considerably since the distributions of the parameter state variables provide the necessary mathematical description of the part parameters. The only schematic or connection information needed is contained in reliability diagrams or, equivalently, state tables. Because of the different techniques involved, the reliability of loaded and of non-loaded systems are discussed separately.

Reliability of non-loaded systems

In Eq. (II.5), the reliability value function for the time interval (t_0, t_1) was defined. In the special case under consideration, only one system parameter, denoted by $A(t)$, need be defined. Using the notation of Eq. (II.3), the range of $A(t)$ is defined in Eq. (13).

$$(13) \quad A(t) = 0 \text{ if } X(t) \in Z_0 \\ = 1 \text{ if } X(t) \in Z_1$$

The tolerance set $S(t)$ for the system parameter $A(t)$ is the single value 0. Then, the reliability for the time interval (t_0, t_1) can be written as follows.

$$R(t_0, t_1) = p(A(\mathcal{T}) = 0, \text{ all } t_0 \leq \mathcal{T} \leq t_1), \text{ where } t_0 < t_1$$

Since repair is not allowed, the fact that $A(t) = 0$ implies that $A(\mathcal{T}) = 0, \text{ all } \mathcal{T} > t$. Thus, the reliability expression

is simplified to that in Eq. (14).

$$\begin{aligned}
 (14) \quad R(t_0, t_1) &= P(A(t_1) = 0 / A(t_0) = 0) \\
 &= \frac{P(A(t_1) = 0, A(t_0) = 0)}{P(A(t_0) = 0)} \\
 &= \frac{P(A(t_1) = 0)}{P(A(t_0) = 0)} = \frac{p_A(0; t_1)}{p_A(0; t_0)}
 \end{aligned}$$

The density function $p_A(0; t)$ is expressed in terms of the system state variable in Eq. (II.4). Each state number for which the system is operational is equivalent to a set of values for the parameter state variables. The binary representation for each such state number exhibits the corresponding values of the parameter state variables. Thus, for the value 0, Eq. (15) is equivalent to Eq. (II.4).

$$(15) \quad p_A(0; t) = \sum_{\text{all } \bar{x}} p_{A/\bar{X}}(0/\bar{x}; t) p_X(\bar{x}; t)$$

In Eq. (15), $p_{\bar{X}}$ is the joint density function of the parameter state variables at time t . This expression can be simplified by evaluating the conditional density terms on the basis of Eq. (13), as shown in Eq. (16).

$$\begin{aligned}
 (16) \quad p_{A/\bar{X}}(0/\bar{x}^*; t) &= 0 \text{ if } \bar{x}^* = Z_1; \text{ i.e., } A(t) = 1 \\
 &= 1 \text{ if } \bar{x}^* \in Z_0, \text{ i.e., } A(t) = 0
 \end{aligned}$$

In Eq. (16), \bar{x}^* is the binary representation of x^* .

By the definition of a non-loaded system in Eq. (6), the joint density function of the parameter state variable can be expressed as the product of marginal distributions. The final

expression for the system parameter's density function is obtained by substituting Eq. (6) and Eq. (16) into Eq. (15), with the result shown in Eq. (17a).

$$(17a) \quad p_A(0;t) = \sum_{Z_0} \prod_{i=1}^k p_{X_i}(x_i;t)$$

The summation in Eq. (17a) extends over all sets of parameter state variable values which are binary representations of state numbers in Z_0 .

$$(17b) \quad p_A(1;t) = 1 - p_A(0;t) = \sum_{Z_1} \prod_{i=1}^k p_{X_i}(x_i;t)$$

In Eq. (17b), the sum extends over all sets of parameter state variable values which are not binary representations of state numbers in Z_0 , even if some of these sets of values have measure zero.

An algorithm for computing either $p_A(0;t)$ or $p_A(1;t)$, called the one-failure algorithm, is now presented. The one-failure algorithm is based on Eq. (17) and exploits the analogy between a reliability diagram and a combinational switching circuit schematic. The following three steps define the one-failure algorithm.

1. Construct a reliability table. The form of the reliability table, which corresponds to a truth table, is the same as that of a state table, except that the heading for the system state variable is replaced in a reliability table by a heading for $A(t)$. All 2^k sets of values for the parameter state

variables are listed, even though some are sets of measure zero.

2. Compute one term of Eq. (17a) from each row of the reliability table for which $A(t) = 0$, or one term of Eq. (17b) from each row for which $A(t) = 1$. If one of these rows is $x'_k \text{ --- } x'_1$, the corresponding term in Eq. (17) is given below.

$$\prod_{i=1}^k p_{X_i}(x'_i; t)$$

In this computation, either all $A(t) = 0$ rows or all $A(t) = 1$ rows are used, not both.

3. Add all terms computed in Step 2. If the $A(t) = 0$ rows were used, the expression for $p_A(0; t)$ is found from Eq. (17a); if the $A(t) = 1$ rows were used, $p_A(1; t)$ is computed as indicated by Eq. (17b).

A general method for simplifying the transmittance expression of a two-terminal combinational switching circuit has been discussed by Caldwell (42, p. 145) and is called the Quine-McCluskey method. The variation of this technique presented below is used to simplify the expression for $p_A(0; t)$ or for $p_A(1; t)$ obtained from the one-failure algorithm.

The Quine-McCluskey method starts with a truth table representing the standard sum form of a transmittance and derives its prime implicants. In the variation presented below, an expression for $p_A(0; t)$ or $p_A(1; t)$ is obtained from the reliability table which is simpler in the sense that it involves fewer terms and fewer factors in some terms than the expression resulting from the one-failure algorithm. This method is

based on the following identity.

$$p_{X_i}(0;t) + p_{X_i}(1;t) = 1, \text{ any } i = 1, \dots, k$$

The analogous expression from Boolean algebra is $Y + Y' = 1$.

The Quine-McCluskey method also used the identity $Y + Y = Y$, for which there is no analogy in terms of density functions, so that the method cannot be applied in toto. This explains why a complete analogy cannot be drawn between a switching circuit schematic and a reliability diagram. That is, in a switching circuit, the transmittances of all possible paths can be added to form the transmittance of the circuit. However, if the reliabilities of all possible paths in a reliability diagram were added, the result would not be the reliability of the system, since all Boolean operations are not valid for a reliability diagram.

The simplification procedure is described by the following rules; Step 1 of the one-failure algorithm is assumed to have been completed.

1. The procedure uses either the binary representations for all rows in the reliability table for which $A(t) = 0$ or those for which $A(t) = 1$. For convenience, the number of 1's in a binary number will be called that number's index. The numbers are separated into groups such that all numbers in any group have the same index.

2. The groups are arranged in a column, beginning with the group of lowest index and continuing with groups of increasing index. A line is drawn between each pair of groups to indicate where the index changes.

3. The numbers in the group with index i are compared to those in the group with index $i+1$. Comparison between groups with indices which differ by more than 1 is not necessary. If two numbers, when compared, differ by one and only one digit, these numbers are checked and combined into a single number with the non-common digit being replaced by a blank. For example, 0100 and 0110 combine to form 01_0, while 0100 and 0010 cannot be combined. Once a number is checked, it is eliminated from consideration. Thus, this technique differs from the Quine-McCluskey method in which an exhaustive comparison is made.

4. After making all possible comparisons, all the numbers containing a blank are recorded in a separate tabulation. Again, the numbers are arranged into groups by their indices and the groups listed as in Step 2.

5. The comparison procedure is then repeated. However, if two numbers are to be combined, their dashes must be in the same position and they must also differ in one and only one bit.

6. A third tabulation is made, the groups arranged, and the comparison and checking procedure repeated. This continues until no new tabulations can be formed.

7. The expression for $p_A(0;t)$ or $p_A(1;t)$ is a sum of terms, one for each unchecked number in all the tabulations. The terms are written as before except that a 1 is written for each blank. For example, 01_1_0 becomes the following, assuming that the columns of the reliability table are labelled X_6, X_5, \dots, X_1 .

$$p_{X_6}(0;t)p_{X_5}(1;t)p_{X_3}(1;t)p_{X_1}(0;t)$$

An example of the simplification procedure is given in terms of the partial reliability table in Fig. 7. Only those lines for which $A(t) = 0$ are listed in Fig. 7. The corresponding reliability diagram is shown in Fig. 3A. In constructing the partial reliability table, it was assumed that all part parameters fail in the open mode. One set of tabulations for the simplification procedure is shown in Fig. 8.

<u>X₆</u>	<u>X₅</u>	<u>X₄</u>	<u>X₃</u>	<u>X₂</u>	<u>X₁</u>	<u>X₆</u>	<u>X₅</u>	<u>X₄</u>	<u>X₃</u>	<u>X₂</u>	<u>X₁</u>
0	0	0	0	0	0	0	0	1	0	1	1
0	0	0	0	0	1	0	0	1	1	0	0
0	0	0	0	1	0	0	0	1	1	0	1
0	0	0	0	1	1	0	1	0	0	0	0
0	0	0	1	0	0	0	1	0	0	0	1
0	0	0	1	0	1	0	1	0	0	1	0
0	0	0	1	1	0	0	1	0	1	0	0
0	0	1	0	0	0	0	1	0	1	0	1
0	0	1	0	0	1	0	1	0	1	1	0
0	0	1	0	1	0						

Fig. 7.--Partial reliability table for Fig. 3A

The required expression is written from the unchecked numbers in Fig. 8 as follows. For simplicity, $p_{X_i}(x;t) = p_i(x)$.

$$\begin{aligned}
 p_A(0) = & p_6(0) \left[p_5(0)p_4(1)p_3(0)p_2(1)p_1(0) + p_5(0)p_4(0)p_2(1)p_1(0) \right. \\
 & + p_5(0)p_3(0)p_2(1)p_1(1) + p_5(1)p_4(0)p_2(0)p_1(1) \\
 & \left. + p_5(1)p_4(0)p_1(0) + p_5(0)p_2(0) \right]
 \end{aligned}$$

<u>000000</u> X	<u>00000_</u> X	<u>000_0_</u> X	<u>00__0_</u>
000001 X	000_10	001_0_ X	
000010 X	00010_ X	<u>010__0</u>	
000100 X	00100_ X		
001000 X	<u>0100_0</u> X		
<u>010000</u> X	00_011		
000011 X	00110_ X		
000101 X	010_01		
000110 X	<u>0101_0</u> X		
001001 X			
001010			
001100 X			
010001 X			
010010 X			
<u>010100</u> X			
001011 X			
001101 X			
010101 X			
<u>010110</u> X			

Fig. 8.--Tabulations of a simplification procedure for Fig. 7

The purpose of the one-failure algorithm and simplification procedure is to provide an almost rote method, which is always applicable, for finding the reliability of a non-loaded system when drift is neglected. As with most universal techniques, short-cut methods exist for certain connection schemes. An example of such a scheme is the series-parallel diagram discussed by Radner (40), Lipp (43), Moskowitz (41), Bazovsky (26) and

others. These authors also use a form of Bayes' theorem which can reduce the work involved in obtaining reliability expressions.

A different approach has been taken by Weiss and Kleinerman (45). This early work studies the set of all paths in a reliability diagram which is drawn with directed lines. If the corresponding "path reliabilities" are combined properly, the system's reliability can be found. This technique seems to have been discarded by later works in the area, probably because of the unnecessary labor involved. This work also lets the reliability of each component be a random variable, which is not in accord with the standard definition.

The analogy between reliability diagrams and switching circuit schematics has been noted by many authors, particularly by Moscovitz (44). This paper utilizes concepts similar to signal flow graphs and equivalent circuits from a non-probabilistic viewpoint. Components are represented by lines and their connections, by nodes. Such a representation is inherently a two-line diagram. Each line is assigned a constant, corresponding to an average reliability over a time interval. A technique is then given for finding an average value for the system's reliability. As Moskowitz recognizes, each line on his diagram must correspond to a different component for his technique to be valid. This rules out an equivalent diagram for non-two-line diagrams such as that in Fig. 3C.

Reliability of loaded systems

Some of the results for non-loaded systems are now extended

to loaded systems. Drift failures are again neglected so that each part parameter can assume only two values. The discussion of loaded systems diverges from that for non-loaded systems after Eq. (14) since, in a loaded system, the joint distribution of the parameter state variables cannot be expressed as the product of marginal distributions.

If the expression for the density function of the system parameter $A(t)$ given in Eq. (II.4) is combined with Eq. (13), Eq. (18) is obtained.

$$(18) \quad p_A(0;t) = \sum_{Z_0} p_X(x;t)$$

The sum in Eq. (18) extends over all values $x \in Z_0$, the set of non-failed states.

As discussed previously, if the lifetimes of all part parameters have exponential distributions of the form shown in Eq. (C.1), the theory of stationary Markov chains can be used to find the density function of the system state variable. Under these conditions, Eq. (11) can be used with Eq. (18) to produce Eq. (19). The reliability of loaded systems is found from Eq. (14) and Eq. (19).

$$(19) \quad p_A(0;t) = \sum_{Z_0} p_{0j}(t)$$

Of course, Eq. (18) is also valid for non-loaded systems. However, the one-failure algorithm does not explicitly use the system state variable and is much easier to apply.

In an interesting paper, Balaban (46) has discussed, among other things, the effects of loading on systems containing active redundancy. This paper is one of the very few which explicitly discusses these effects. However, this study applies only to two elements in parallel, both of which have exponential failure distributions. Although formulas in terms of general failure densities are given, the failure process must have the stationary property before the formulas are valid. Only exponentially distributed lifetime variables possess this property.

Some simple reliability diagrams with exponentially failing components are also discussed by Bazovsky (26, p. 134) in which loads are shared by the components.

Two catastrophic failure modes

The results of the previous section are now extended to systems in which part parameters can fail catastrophically in either of two modes. In this section, the set of catastrophic values $C_1(t)$ for at least one part parameter in a system is taken as a two point set. However, a part parameter cannot take on both catastrophic values during one period of operation. That is, if $B_1(t) = c_{1j}$, then $B_1(\tau) = c_{1j}$, all $\tau \geq t$ for $j = 1, 2$.

This section is subdivided in the same manner as was the previous section.

Since two failure modes are distinguished for part parameters, two catastrophic failure modes are also distinguished for the system. For convenience, these modes are called the open

and short modes of system failure.

The range of the system state variable

As before, the range of a system state variable is defined in terms of parameter state variables. If $r = 3$ in Eq. (1), and if Eq. (20) is satisfied, $M_k M_{k-1} \dots M_1$ is the base three representation of N_{10} .

$$(20) \quad N_{10} = \sum_{i=1}^k M_i 3^{i-1}, \quad 0 \leq M_i \leq 2.$$

Assuming that all part parameters are operational at $t = 0$, when the system is turned on, state numbers are assigned by the following rules.

1. If, at time t , the system has not failed, let $M_i = X_i(t)$. Let the number N_{10} obtained from Eq. (20) be the state number at t . The set of all such numbers is denoted by Z_0 .

2. Let y_1, \dots, y_j be all numbers obtained as in Rule 1, for which the system is in the failed-open state. Then $Z_{11} = \min. (y_1, \dots, y_j)$ is the state number when the system has failed in the open mode.

3. If the word "open" is replaced by "short" in Rule 2, then $Z_{12} = \min. (y_1, \dots, y_j)$ is the state number when the system has failed in the short mode.

The set of values, (Z_{11}, Z_{12}) is denoted by Z_1 .

The base three representation of a state number in Z_0 , when written with k digits, exhibits the values of the parameter state variables.

All possible states of a system can be conveniently displayed in the form of a state table. The only change from the state table discussed in the previous section is the base in which the state numbers are written. This information can also be recorded schematically on a reliability diagram which is interpreted in the same manner as was discussed previously. An example is given in Fig. 9.

The distribution of the system state variable

Lifetime random variables are discussed first, followed by techniques for obtaining the required distributions.

Lifetime random variables

The possibility of three values for each parameter state variable implies that the distribution of $X_i(t)$ cannot be specified in terms of L_i alone. The mode of failure is also important. This additional information is included by defining the two functions, $K_{i1}(t)$ and $K_{i2}(t)$ in Eq. (21).

$$(21) \quad K_{i1}(t) = P(B_i(t) = c_{i1} / L_i \leq t);$$

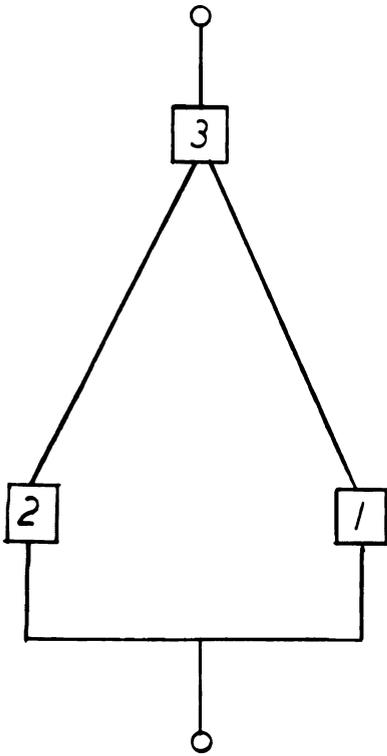
$$K_{i2}(t) = P(B_i(t) = c_{i2} / L_i \leq t)$$

If $L_i \leq t$, the i^{th} part parameter must have failed in one of the two modes so that the following conditions must be satisfied.

$$K_{i1}(t) + K_{i2}(t) \equiv 1; \quad K_{ij}(t) \geq 0, \quad \text{all } t, \quad j = 1, 2.$$

The relationship between a parameter state variable, lifetime variable, and values of a part parameter are given in Eq. (22) in which the (continuous) density function of L_i is denoted by f_i .

$x_3(t)$	$x_2(t)$	$x_1(t)$	$x(t)$
0	0	0	0
0	0	1	1
0	0	2	2
0	1	0	3
0	1	1	4
0	1	2	5
0	2	0	6
0	2	1	7
0	2	2	8
1	0	0	4
1	0	1	4
1	0	2	4
1	1	0	4
1	1	1	4
1	1	2	4
1	2	0	4
1	2	1	4
1	2	2	4
2	0	0	18
2	0	1	19
2	0	2	20
2	1	0	21
2	1	1	4
2	1	2	20
2	2	0	20
2	2	1	20



$$Z_{11} = 4$$

$$Z_{12} = 20$$

Fig. 9.--Example of a state table and corresponding reliability diagram.

$$(22a) \quad p_{X_1}(0;t) = P(B_1(\mathcal{T}) \in D_1(\mathcal{T}), 0 \leq \mathcal{T} \leq t)$$

$$P(L_1 > t) = \int_t^{\infty} f_1(z) dz$$

$$(22b) \quad p_{X_1}(1;t) = P(B_1(t) = c_{i1} / L_{i1} \leq t) P(L_1 \leq t)$$

$$= K_{i1}(t) \int_0^t f_1(z) dz$$

$$(22c) \quad p_{X_1}(2;t) = P(B_1(t) = c_{i2} / L_i \leq t) P(L_1 \leq t)$$

$$= K_{i2}(t) \int_0^t f_1(z) dz$$

As an example, the distribution of a parameter state variable when the lifetime random variable has the exponential distribution in Eq. (9) will be found.

The relative frequency of short and open failure are assumed to remain constant with time so that the total failure rate, H_1 , in Eq. (9) is the sum of two constant failure values.

$$H_1 = H_{i1} + H_{i2}$$

The 1 and 2 subscripts imply open and short failure rates, respectively. Then, the K functions are shown below.

$$K_{i1}(t) = H_{i1} / H_1 \quad ; \quad K_{i2} = H_{i2} / H_1$$

Under these conditions, the density function in Eq. (22) can be written as in Eq. (23)¹.

¹Care must be taken when combining two failure rates. A conceptual error, such as that made by Bazovsky (26, p.138) in his discussion of switch failures, can easily creep into a development. Bazovsky lets a switch have a constant fail-open rate H_0 and a constant fail-short rate H_s . He then makes the following statements.

$$q_0 = \text{Probability of failing open} = 1 - \exp(-H_0 t)$$

$$(23a) p_{X_i}(0;t) = \exp(-H_i t)$$

$$(23b) p_{X_i}(1;t) = (H_{i1} / H_i) [1 - \exp(-H_i t)]$$

$$(23c) p_{X_i}(2;t) = (H_{i2} / H_i) [1 - \exp(-H_i t)]$$

The state-changing process

The state-changing process is essentially the same as that discussed in previous section. The system leaves state i , $i \in Z_0$ if a 0 digit in the base three representation of i changes to a 1 or to a 2, which implies that the corresponding part parameter has failed. The system cannot leave either state in Z_1 , nor can it return to any state. Transition charts are drawn in the same manner as in the previous subsection.

Distribution of the system state variable for non-loaded systems

For state numbers in Z_0 , the density function of the system state variable can be found from Eq. (6) by using Eq. (22) for the separate factors. The value of this density function for the two states in Z_1 is not as easy to compute. The sum of these two terms must satisfy Eq. (24)

$$(24) \sum_{Z_1} p_X(x;t) = 1 - \sum_{Z_0} \prod_{i=1}^k p_X(x_i;t)$$

$$q_s = \text{Probability of failing short} = 1 - \exp(-H_s t)$$

$$r = \text{Probability of no failure} = 1 - (q_0 + q_s) = \exp(-H_0 t)$$

+ $\exp(-H_s t) - 1$

Now, if $t > \min.(\ln 2/H_0, \ln 2/H_s)$, then $r < 0$, which, of course, violates the definition of the probability of an event. The error here is in the assumption that the two possible failure events are independent and mutually exclusive.

One is tempted to compute the separate terms on the left of Eq. (24) from formulas similar to Eq. (6), using those rows in the state table for which $X(t) = Z_{11}$ or Z_{12} . However, the sum of the two terms computed in this manner might not satisfy Eq. (24). If all 3^k possible rows were entered on the state table, regardless of whether some are sets of measure zero, Eq. (24) would be satisfied, but there is no assurance that the values of the density function so obtained would correspond to the actual failure probabilities.

At present, the density function's values at the shut down states can be found only when each parameter state variable has the distribution given in Eq. (23). This situation is discussed with loaded systems.

Of course, for simple cases, such as all blocks in the reliability diagram in series and parallel, the density function can always be computed.

Distribution of the system state variable for loaded systems

As in the previous section, only a special case is discussed, namely when the state variable of each part parameter has the distribution given in Eq. (23). The failure rates H in this distribution may change when the system changes states. Again, the theory of continuous parameter, stationary Markov chains is applied directly to this problem. The initial condition of the process are given by Eq. (10). The density function can be evaluated from Eq. (11) by either of the two techniques discussed previously. The relationships between the q 's, which determine the transition probabilities, and the lifetime

distributions are shown below.

The density function f_{ij} of the lifetime variable associated with the j^{th} part parameter when the system is in state i is assumed to be of the form shown in Eq. (25).

$$(25) \quad \begin{aligned} f_{ij}(z) &= 0, \quad z < 0 \text{ or } i \in Z_1 \\ &= 0 \text{ if } X_j(t) \neq 0 \text{ and } i \in Z_0 \\ &= H_{ij} \exp(-H_{ij}z) \text{ otherwise.} \end{aligned}$$

In Eq. (25), $H_{ij} = H_{ij1} + H_{ij2}$, where H_{ij1} is the open failure rate and H_{ij2} , the short failure rate.

If $X_{r_m}(t)$, $m = 1, 2, \dots, n$ are all 0 when the system is in state i , then q_i is given by Eq. (26)

$$(26) \quad \begin{aligned} q_i &= \sum_{m=1}^n H_{ir_m} \text{ if } i \in Z_0 \\ &= 0 \quad \text{if } i \in Z_1. \end{aligned}$$

If the failure of $B_r(t)$ causes the system to go from state i to state j , $i, j \in Z_0$, q_{ij} is one of two failure rates.

If $X_r(t) = 1$ in the binary representation of j , q_{ij} is given by Eq. (27a).

$$(27a) \quad q_{ij} = H_{ir1}$$

If $X_r(t) = 2$ in the binary representation of j , q_{ij} is given by Eq. (27b).

$$(27b) \quad q_{ij} = H_{ir2}$$

If $i \in Z_0$, $j = Z_{11}$ and if the open failure of any of the part parameters $B_r(t)$, \dots , $B_v(t)$ causes the system to go from i to j in one step, q_{ij} is given by Eq. (28a).

$$(28a) \quad q_{ij} = \sum_{m=1}^v H_{ir_m} 1.$$

If $i \in Z_0$, $j = Z_{12}$ and if the short failure of any of the part parameters $B_{r_1}(t), \dots, B_{r_h}(t)$ causes the system to go from i to j in one step, q_{ij} is given by Eq. (28b).

$$(28b) \quad q_{ij} = \sum_{m=1}^h H_{ir_m} 2$$

Reliability

Special techniques for finding reliability expressions when drift is neglected are now presented. For notational simplicity, all part parameters are allowed to have both failure modes. Non-loaded and loaded systems are discussed separately.

Reliability of non-loaded systems

Since drift failures are neglected, the system parameter $A(t)$ defined in Eq. (13) is again the only system parameter of interest. Only two values of $A(t)$ need be defined, even though the set Z_1 in Eq. (13) contains two points, since the reliability value function does not distinguish the modes of system failure.

The expression for reliability in Eq. (14), that for the density function of $A(t)$ in Eq. (17), and the one-failure algorithm are all directly applicable to this case. The only difference is that in the previous section, base 2 numbers are used in the reliability table and in this section, base 3 numbers are used.

The simplification procedure discussed previously would

have to be extended to cover this case. Three , rather than two, base three representations have to be combined to eliminate a factor.

The one-failure algorithm, although applicable here, requires 3^k rows in the reliability table, where k is the number of part parameters. A second algorithm, called the two-failure algorithm, is now presented which requires only 2^{k+1} rows and allows direct utilization of the simplification procedure. The two-failure algorithm is described by the following three steps. Its theoretical basis is given in Appendix D.

1. Assume that no part parameter can fail in the short mode. Find an expression for either of two functions, called $p^{(1)}(0;t)$ and $p^{(1)}(1;t)$ by using the one-failure algorithm and simplification procedure. Only 0's and 1's are listed in the reliability table. The bits in the unchecked numbers resulting from the simplification procedure are interpreted as follows.

- i) 1 becomes $p_{X_i}(1;t)$
- ii) 0 becomes $1 - p_{X_i}(1;t)$, not $p_{X_i}(0;t)$

2. Assume that no part parameter can fail in the open mode. Find an expression for either of two functions, called $p^{(2)}(0;t)$ and $p^{(2)}(1;t)$ by using the one-failure algorithm and simplification procedure. The reliability table contains only 0's and 2's. The results of the simplification procedure are interpreted as follows.

- i) 2 becomes $p_{X_i}(2;t)$
- ii) 0 becomes $1 - p_{X_i}(2;t)$, not $p_{X_i}(0;t)$

3. Determine $p_A(0;t)$ from one of the following relations.

$$p_A(0;t) = p^{(1)}(0;t) + p^{(2)}(0;t) - 1$$

$$p_A(0;t) = 1 - p_A(1;t) = 1 - [p^{(1)}(1;t) + p^{(2)}(1;t)]$$

A method similar to the two-failure algorithm has been used by Price (47), without proof, to compute the reliability of parallel components.

The situation when all blocks in a reliability diagram have identical failure distributions has been studied by Barlow and Hunter (34). The diagram studied is a series and parallel connection of subdiagrams, each of which is constructed from blocks connected in series and parallel. In this case, the reliability can be written directly, using a multinomial distribution.

As an example, an expression for $p_A(0;t)$ will be found for the system whose reliability diagram is given in Fig. 10, using the two-failure algorithm.

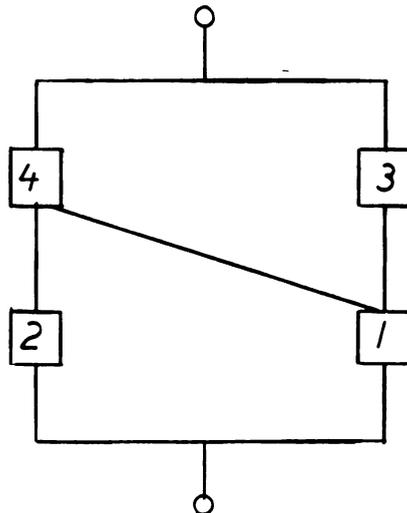


Fig. 10.--Example of a reliability diagram

The details of Step 1 of the two-failure algorithm are shown in Fig. 11. Only those rows for which $A(t) = 1$ are listed in the reliability table. The following expression

results from this step. For convenience, $p_{X_i}(x;t)$ is written as $p_i(x)$.

$$\begin{aligned} p^{(1)}(1;t) &= [1 - p_4(1)] p_2(1) p_1(1) + \\ & p_4(1) [(1 - p_3(1)) p_1(1) + p_3(1)] \\ &= p_2(1) p_1(1) + \\ & p_4(1) [p_1(1) + p_3(1) - p_1(1) p_3(1) - p_2(1) p_1(1)] \end{aligned}$$

X_4	X_3	X_2	X_1	A			
0	0	1	1	1	0011 X	0_11	11__
0	1	1	1	1	1001 X	10_1	
1	0	0	1	1	<u>1100 X</u>	<u>110_ X</u>	
1	0	1	1	1	0111 X	111_ X	
1	1	0	0	1	1011 X		
1	1	0	1	1	1101 X		
1	1	1	0	1	<u>1110 X</u>		
1	1	1	1	1	1111 X		

Fig. 11.--Example of Step 1 of the two-failure algorithm

The details of Step 2 are shown in Fig. 12. Only those rows for which $A(t) = 0$ are listed in the reliability table. The following expression is obtained.

$$\begin{aligned} p^{(2)}(0;t) &= [1 - p_4(2)] [p_3(2)(1 - p_1(2)) + 1 - p_3(2)] \\ &+ p_4(2) [1 - p_2(2)] [1 - p_1(2)] \\ &= 1 - p_3(2) p_1(2) - \\ & p_4(2) [p_1(2) + p_2(2) - p_1(2) p_2(2) - p_1(2) p_3(2)] \end{aligned}$$

<u>X₄</u>	<u>X₃</u>	<u>X₂</u>	<u>X₁</u>	<u>A</u>			
0	0	0	0	0	<u>0000 X</u>	<u>000_ X</u>	00__
0	0	0	2	0	0002 X	002_ X	
0	0	2	0	0	0020 X	02_0	
0	0	2	2	0	0200 X	2_00	
0	2	0	0	0	<u>2000 X</u>		
0	2	2	0	0	0022 X		
2	0	0	0	0	0220 X		
2	2	0	0	0	<u>2200 X</u>		

Fig. 12.--Example of Step 2 of the two-failure algorithm

The results of Step 3 are obtained as follows.

$$\begin{aligned}
 p_A(0;t) &= 1 - p^{(1)}(1;t) + p^{(2)}(0;t) - 1 \\
 &= p^{(2)}(0;t) - p^{(1)}(1;t).
 \end{aligned}$$

Loaded systems

If the lifetime variables of all part parameters have the distribution shown in Eq. (25), the theory of Stationary Markov chains can be used to find $p_A(0;t)$ in Eq. (14). Equation (19) is again used for this purpose, on the basis of Eq. (25) through Eq. (28).

CHAPTER IV

DRIFT EFFECTS

Two types of drift in redundant systems are discussed in this chapter, neither of which has been treated in the literature. In the first type, the drift characteristics of each non-failed part parameter depend on the length of time the system operates; in the second, the drift characteristics depend both on the states the system assumes and on the length of time the system remains in each state. Both types are applicable to loaded and to non-loaded systems.

The first section in this chapter outlines the particular drift problem to be discussed. Methods for finding the joint distribution of the system parameters at a given time are presented in the second section. These methods provide a basis for the third section which deals with time variations in the distributions of part and system parameters.

The drift problem

If the joint distribution of the system parameters is known at any time, the reliability and utility value functions can be found. However, for definiteness, only the measure of utility defined by Eq. (1a) will be treated in this chapter. The techniques presented can be extended to cover other value functions.

$$(1a) \quad Q(t) = P(A_1(t) \in S_1(t), \text{ all } i=1, \dots, n)$$

$$= \int_{\bar{S}(t)} f_{\bar{A}}(\bar{a}; t) d\bar{a}$$

The function $Q(t)$, sometimes called pointwise reliability, is the probability that the system is operating satisfactorily at time t . The notation $\bar{S}(t)$ represents the tolerance sets for the system parameters $\bar{A}(t)$.

In order to use the General Model to find expressions for $Q(t)$, Eq. (II.4) is substituted into Eq. (1a) with the result shown in Eq. (1b).

$$(1b) \quad Q(t) = \sum_{\mathbf{x}} p_X(\mathbf{x}; t) \int_{\bar{S}(t)} f_{\bar{A}/X}(\bar{a}/\mathbf{x}; t) d\bar{a}$$

If $\mathbf{x} \in Z_1$, i.e., if the system is in a failed state, the corresponding integral in Eq. (1b) is assumed to be zero for any tolerance sets. Since, by definition, the system is operationally useless after it has failed, this is not a restrictive assumption. Equation (1b) then becomes Eq. (1c).

$$(1c) \quad Q(t) = \sum_{Z_0} p_X(\mathbf{x}; t) \int_{\bar{S}(t)} f_{\bar{A}/X}(\bar{a}/\mathbf{x}; t) d\bar{a}$$

The sum in Eq. (1c) extends over all $\mathbf{x} \in Z_0$, the set of non-failed states.

When the system has not failed, those part parameters which have not failed and which have not been removed from the system because of the failure of other part parameters are called the operational part parameters. When the system is in

state $x \in Z_0$, the set of operational part parameters is denoted by $\bar{B}_x(t)$. The members of this set change when the system changes states as do the forms of the system parameters. Thus, in order to evaluate the integrals in Eq. (1c), the joint distribution of the system parameters for every state in Z_0 must be found at time t . These joint distributions can be derived by the techniques discussed in the next section if the joint distribution of the operational part parameters for every state in Z_0 are known at time t . Methods for deriving these joint distributions from drift data are presented in the last section.

Distributions of system parameters at one time

This section is devoted to a review of the available techniques for studying the joint distribution of the system parameters at a particular time. Specifically, the distribution of the A's in Eq. (2) will be studied, where $\bar{B} = (B_1, \dots, B_k)$ represents a set of operational part parameters.

$$(2) \quad A_i = v_i(B_1, \dots, B_k) \quad i=1, \dots, n$$

The first two subsections present techniques for finding the entire distributions of functions of random variables, the first using analytical methods and the second, Monte Carlo techniques. The methods available for finding certain measures of distributions are taken up in the third subsection. Some special techniques for the important special case of linear functions of independent random variables are presented in the last subsection.

Analytical techniques for finding complete distributions

The exact form of the joint distribution of the B's in Eq. (2) is assumed known; i.e., the density or cumulative distribution function describing this joint distribution is given in closed, analytical form. The v_i functions are also assumed to be known.

The first technique to be discussed has been presented by many authors, e.g., Cramer (48, p. 292). This technique is first presented for the case when $n = k$ in Eq. (2) and applies only when all B's are continuous random variables. The following two conditions are assumed to be satisfied for all (vector) values \bar{b} of the set of random variables $\bar{B} = (B_1, \dots, B_k)$ for which the joint density function of \bar{B} is non-zero.

A) The functions v_i are everywhere unique and continuous and have continuous partial derivatives $\partial v_i / \partial z_j$ $i, j = 1, \dots, k$.

B) The relationships in Eq. (2) define a one-to-one correspondence between the points \bar{b} and \bar{a} , which are values of \bar{B} and \bar{A} , so that the inverse relationships written as Eq. (3) exist, where the w_i are unique.

$$(3) \quad B_i = w_i(A_1, \dots, A_k) \quad i = 1, \dots, k=n.$$

The functions in Eq. (2) and inverses in Eq. (3) are written in terms of values of the random variables in Eq. (4).

$$(4a) \quad a_i = v_i(b_1, \dots, b_k) \quad i = 1, \dots, k=n$$

$$(4b) \quad b_i = w_i(a_1, \dots, a_k)$$

The probability element of the joint distribution function of the A's is expressed in terms of that for the B's in Eq. (5).

$$(5) \quad f_{\bar{A}}(\bar{a})d\bar{a} = |J| f_{\bar{B}}(\bar{b})d\bar{b}$$

The values \bar{b} in the right-hand side of Eq. (5) must be replaced by values \bar{w} from the inverse transformation in Eq. (4b). The function $|J|$ is the absolute value of the Jacobian of the transformation defined by Eq. (4a).

$$J = \frac{\partial(w_1, \dots, w_n)}{\partial(a_1, \dots, a_n)}$$

The Jacobian can also be expressed as follows, where, in the result, Eq. (4b) is used.

$$J = \left[\frac{\partial(v_1, \dots, v_n)}{\partial(b_1, \dots, b_n)} \right]^{-1}$$

For the special case when $n=1$, the transformation in Eq. (4a) is written in Eq. (6a).

$$(6a) \quad a_1 = v_1(b_1, \dots, b_k) \\ a_i = b_i, \quad i = 2, \dots, k.$$

If conditions A and B are satisfied, the unique inverse transformation given in Eq. (6b) exists.

$$(6b) \quad b_1 = w_1(a_1, \dots, a_k) \\ b_i = a_i, \quad i = 2, \dots, k.$$

The Jacobian of this transformation becomes the following.

$$J = \frac{\partial w_1(a_1, \dots, a_k)}{\partial a_1} = \left[\frac{\partial v_1(b_1, a_2, \dots, a_k)}{\partial b_1} \right]^{-1}$$

In this expression, b_1 is replaced, after differentiation, by $w_1(a_1, \dots, a_k)$. Then, Eq. (5) becomes the following.

$$f_{\bar{A}}(\bar{a})d\bar{a} = |J| f_{B_1, B_2, \dots, B_k}(w_1(a_1, \dots, a_k), a_2, \dots, a_k)d\bar{a}$$

The (marginal) distribution of A_1 is then found by integration as shown in Eq. (7).

$$(7) \quad f_{A_1}(a_1) = \int_{a_n} \dots \int_{a_2} f_{\bar{A}}(\bar{a}) da_2 \dots da_n$$

The extension to the case when $n > 1$, but $n \neq k$ is obvious.

A second technique for finding the joint distribution of the A's in Eq. (2) has been discussed by Middleton (49, p. 21). The characteristic function of the vector variable \bar{A} is expressed in terms of that for \bar{B} . The joint density function of \bar{A} is then found by an inverse Fourier transform.

In almost all applications of the above techniques, the resulting density functions are of non-standard form, and are not tabulated. However, in a few special cases, they are of recognizable form. For example, the sum of independent, normally distributed random variables is itself normally distributed. The chi-squared, student's t and F distributions were all derived from certain combinations of random variables. However, these special forms have only a limited application to the present problem.

Monte Carlo techniques for finding complete distributions

As applied to the problem at hand, "Monte Carlo" is a synonym for "sampling by computer". A method using this approach for approximating the density function of a known function of random variables has been presented by Hellerman and Racite (17). Their procedure, called "synthetic sampling", approximates the density function of A in Eq. (8).

$$(8) \quad A = v(B_1, \dots, B_k)$$

The cumulative distribution function for each of the random

variables B_1, \dots, B_k is assumed known. Each B is a continuous random variable so that there exists a unique inverse of these distribution functions for any number between 0 and 1. A random number generator¹ is assumed to be programmed in the digital computer.

The first computational step is obtaining k numbers r_1, \dots, r_n from the random number generator. Each of these is used to compute one inverse of a cumulative distribution function so that k numbers b_1, \dots, b_k are found, where $b_i = F_{B_i}^{-1}(r_i)$. That is, a random sample from the distribution of each B_i is obtained.

One value of A is then computed from the b 's using Eq. (8). The whole process is repeated as often as is necessary until a number of values of A are obtained. Each is interpreted as a random sample from the distribution of A . These samples are then grouped in intervals on the axis representing values of A . The proportion of numbers in each interval is plotted in histogram fashion to approximate the density function of A .

Some generalizations can be made on the technique described above. The B 's need not be continuous random variables. If one is discrete, its cumulative distribution function is a step function with jumps at each point for which the density function is non-zero. A one-to-one correspondence between the saltus at each jump and a subinterval of the $[0,1]$ interval can

¹A random number generator is a set of computations which will produce random samples from a distribution that is uniform on the closed interval $[0,1]$. They have been discussed, for example, by Tausky and Todd (50).

be made. Thus, these subintervals can be related to mass points of the distribution.

Obviously, if the distribution of more than one function of the B's is desired, each set of numbers (r_1, \dots, r_k) is used to compute a value for all functions of interest.

Techniques for finding measures of a distribution

In situations where the complete distribution of a function is not easily found, certain measures of the distribution can be useful in a failure analysis. Some of these measures and methods for estimating them are discussed in this subsection.

Moments

The moments, μ_1, μ_2, \dots of a distribution having cumulative distribution F are defined in Eq. (9).

$$(9) \quad \mu_j = \int_{-\infty}^{\infty} x^j dF(x)$$

If $\mu_0 \equiv 1$, and if the moments μ_1, μ_2, \dots corresponding to a distribution function $F(x)$ are all assumed to be finite, and if the series shown below is absolutely convergent for some $r > 0$, then $F(x)$ is the only distribution function having these moments.

$$\sum_{j=0}^{\infty} (\mu_j / j!) r^j$$

If $C(s)$ is the characteristic function for the distribution defined by $F(x)$, then $C(s)$ is given by the following, where i is the complex operator.

$$C(s) = \sum_{j=0}^{\infty} (\mu_j/j!)(is)^j$$

These relationships between the moments, distribution function, and characteristic function were pointed out by Cramer (48, p. 174).

Tchebycheff's inequality

It is seldom easy to find or estimate all the moments of a distribution. Fortunately, many reliability computations require only an estimate of the percentage of the density function which is within certain limits. Such estimates can be made from approximations to the mean and variance of a distribution by means of the Tchebycheff inequality given in Eq. (10), which has been discussed by Cramer (48, p. 183).

$$(10) \quad P(|Y - m| \geq ks) \leq (1/k^2), \quad k > 0$$

In Eq. (10), Y is a random variable having (finite) mean m and variance s^2 . Thus, the probability that the value of a random variable lies in the range $m \pm 4.5s$ is at least 0.95, no matter what the distribution of the random variable, as long as it has finite mean and variance.

The estimates provided by the inequality in Eq. (10) is, in many cases, quite conservative. For example, if Y has a normal distribution, the following is true.

$$P(|Y - m| \geq 3s) \doteq 0.0044$$

However, Eq. (10) produces an estimate of 0.111. Nevertheless, Eq. (10) provides one of the only analytical means for using estimates of m and s to place bounds on probabilities.

Estimation of the mean and variance

Methods for estimating the mean, m and variance, s^2 , for a random variable A of the form shown in Eq. (8) are now considered.

The first method to be discussed is called the "propagation of error" method. A very thorough review of this technique has been given by Murphy (51) in which the relationships between cumulants of A and the B 's are studied. A simplified form has been presented by Heyne (52) for functions of two variables.

The following results are derived in Appendix E.

The mean of A is estimated from Eq. (E.5), in which

$$m_i = E(B_i).$$

$$(E.5) \quad m \doteq v(m_1, \dots, m_k)$$

The variance of A is estimated from Eq. (E.6), where s_r^2 is the variance of B_r and v_r is the partial derivative of v with respect to B_r , evaluated at the means of the B 's.

$$(E.6) \quad s^2 \doteq \sum_{r=1}^k v_r^2(m_1, \dots, m_k) s_r^2$$

Equation (E.6) rests on the following three assumptions.

- 1) Equation (E.5) is satisfied.
- 2) The B 's are statistically independent.
- 3) The higher order terms in Eq. (E.1) are neglected.

If condition 2 is not satisfied, the covariance terms in Eq. (E.4) must be added. If condition 3 is not satisfied, covariance terms resulting from the higher order terms in the Taylor expansion in Eq. (E.1) must also be added. If condition 1

is not satisfied, Eq. (E.6) is essentially useless. In this case, one must resort to numerical integration of the expression for the mean and second moment, since for any random variable A, the variance, if it exists, is given by the following expression.

$$s^2 = \mu_2 - m^2$$

The discussion of a Monte Carlo technique in the previous subsection suggests another possible approach to the problem of approximating m and s^2 . Since random samples of A are obtained in the Monte Carlo technique, methods of statistical inference could be used for estimating m and s^2 . A review of such techniques is beyond the scope of this thesis.

Linear functions of independent random variables

The very important special case of linear functions of independent random variables is now discussed. Of course, all techniques described previously also apply here, but some special techniques are of interest. The importance of this case stems from the fact that non-linear functions can sometimes be approximated by linear functions, using Eq. (E.1) as shown in Eq. (11).

$$(11) \quad v(B_1, \dots, B_k) \doteq J_0(\bar{m}) + \sum_{r=1}^k v_r(\bar{m})B_r$$

The following notation is used in Eq. (11).

$$J_0(\bar{m}) = v(m_1, \dots, m_k) - \sum_{r=1}^k m_r v_r(m_1, \dots, m_k)$$

$$v_r(\bar{m}) = v_r(m_1, \dots, m_k)$$

The $v_r(\bar{m})$ factors are sometimes called sensitivity terms since they can also be obtained from the differential of v .

For convenience, the function shown in Eq. (12), rather than that in Eq. (11), will be discussed. In Eq. (12), the ζ_i are mutually independent random variables.

$$(12) \quad A = \sum_{i=1}^k \zeta_i$$

Two analytical techniques for finding the distribution of A have been presented by Cramer (48, p. 158). The first is concerned with characteristic functions. If $C_i(s)$ is the characteristic function of ζ_i and $C_A(s)$ that of A , then they are related by Eq. (13).

$$(13) \quad C_A(s) = \prod_{i=1}^k C_i(s)$$

This method has been used, for example, by Depian and Grisamore (22).

The second method uses the convolution integral, shown symbolically in Eq. (14).

$$(14) \quad F_A(x) = F_1(x) * F_2(x) * \dots * F(x)$$

In Eq. (14), $F_i(x)$ is the cumulative distribution function of

i . The star symbols are interpreted as follows.

$$F_1(x) * F_2(x) = \int_{-\infty}^{\infty} F_1(x-z) dF_2(z) = F_2(x) * F_1(x)$$

It must be emphasized that Eq. (13) and Eq. (14) are valid

only if the ζ_i are independent random variables, related to A by Eq. (12).

An interesting method for approximating the distribution of A in Eq. (12) has been presented by Gray (53) for the case when all ζ_i are continuous random variables. Piecewise linear approximations to the density functions of the ζ 's are made and an algorithm for approximating the density function of A, section by section, is given.

Time variations

Methods for finding the joint distribution of a set of operational part parameters, knowing their drift characteristics are presented in this section. If this joint distribution is derived for every non-failed state of the system at time t, the techniques of the previous section can be applied to find the joint distribution of the system parameters, and Q(t) can be computed from Eq. (1c).

The drift phenomenon is described qualitatively in the first subsection. The second and third subsection treat independent and dependent drift, respectively.

The drift phenomenon in a part parameter

As outlined by Golubjatnikov (54), the variations in the value of a part parameter due to drift can be segregated into three categories, viz, initial variations, short term drift, and long term drift.

Initial variations account for differences in the values of several new part parameters from the same lot. Such variations are caused by the manufacturing process and are usually centered

about a rated, or nominal value. A normal distribution is often used to describe this variation with the mean as the nominal value.

Short and long term drift describe non-catastrophic changes in the value of a part parameter in time. Long term drift is a gradual deterioration resulting in eventual failure. The limiting or end-of-life value depends on the particular system and on the tolerance sets assigned to system parameters. Short term drift, on the other hand, is a random fluctuation about some norm, caused by random changes in the environment. This norm includes the effects of long term drift. The percentage of variation caused by short term drift is usually smaller than that caused by long term drift.

Independent drift

Independent drift refers to the situation when the drift characteristics of the operational part parameters, at any time, depend only on the length of time the system has operated. That is, the environment which causes the drift does not change as long as the system is operating.

The sets of operational part parameters are, at any time, assumed to be statistically independent random variables. If $\bar{B}_x(t) = (B_{r_1}(t), \dots, B_{r_h}(t))$ is a set of operational part parameters, their joint density function, for all t such that $X(t) = x$, is shown in Eq. (15).

$$(15) \quad f_{\bar{B}_x}(\bar{b}; t) = \prod_{i=1}^h f_{B_{r_i}}(b; t)$$

The drift characteristics for a single operational part parameter $B_i(t)$ will now be discussed. These characteristics are expressed as the sum of two random variables in Eq. (16).

$$(16) \quad B_i(t) = B_i(0) + B_i^{(d)}(t)$$

The system is assumed to be turned on at $t = 0$ so that, in Eq. (16), $B_i(0)$ is a random variable representing initial variation. The random variable $B_i^{(d)}(t)$ represents the drift of $B_i(t)$ from its initial value at the end of a time interval of length t and includes both short and long term drift. For any time t for which $B_i(t)$ is operational, $B_i(0)$ and $B_i^{(d)}(t)$ are assumed to be statistically independent, an assumption that has been made by Xavier et. al. (55).

Two methods for determining the distributions of the separate terms in Eq. (16) are now considered.

Empirical method

The first method relies upon measured data and is used to find the distribution of $B_i(t)$ at one particular time. This method has been suggested by Xavier et. al. (55) and is now quantified in terms of the General Model

To find the distribution of $B_i(0)$, a group of components from the same lot are placed on test simultaneously. The value of $B_i(0)$ is measured on all components and the data is plotted in histogram fashion as shown in Fig. 13.

The ordinate of Fig.13 represents the percentage of values of $B_i(0)$ falling within the various intervals. A (discrete) density function can be generated from the histogram as follows. The possible values of $B_i(0)$ are denoted by b_{ij} .

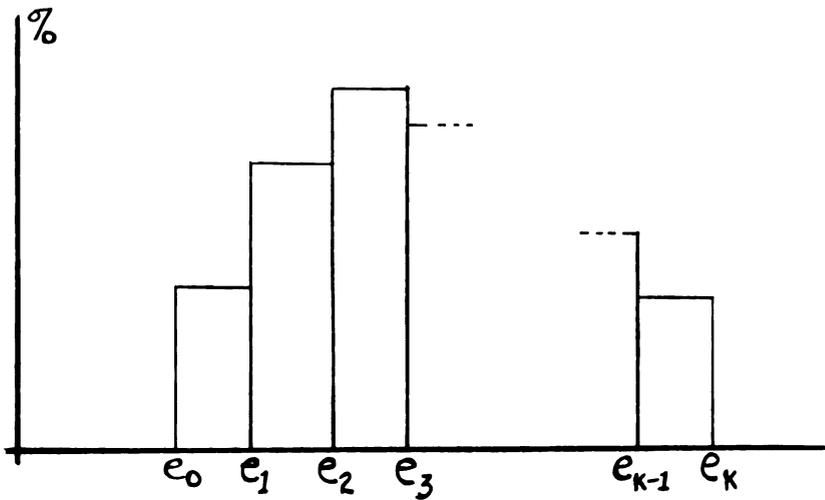


Fig. 13.--Histogram of measured data

$$b_{ij} = (\frac{1}{2})(e_j + e_{j-1}) \quad , \quad j=1, \dots, K$$

If p_{ij} is the ordinate of the histogram for the interval (e_j, e_{j-1}) , then the density function for $B_i(0)$ is given in Eq. (17).

$$(17) \quad p_{B_i}(b_i; 0) = p_{ij} \quad \text{if } b_i = b_{ij}, \quad j=1, \dots, K \\ = 0 \quad \text{otherwise}$$

The distribution of the change from the initial value at time t , $B_i^{(d)}(t)$, is often known from specifications. If it is not, the above technique can be used to measure it. In any event, the (discrete) density function for this distribution is shown in Eq. (18).

$$(18) \quad p_{B_i}^{(d)}(b_i; t) = p_{ij}^{(d)} \quad \text{if } b_i = b_{ij}^{(d)} \quad j=1, \dots, k \\ = 0 \quad \text{otherwise}$$

The distribution of $B_i(t)$ at time t is found by using Eq. (16), Eq. (17), and Eq. (18) with Eq. (14). The resulting density function is shown in Eq. (19).

$$(19) \quad p_{B_i}(b_i; t) = \sum_{j=1}^K p_{B_i}^{(d)}(b_i - b_{ij}; t) p_{ij}$$

$$\quad \text{if } b_i = b_{ij} + b_{ir}^{(d)}, \quad j=1, \dots, K; \quad r=1, \dots, k$$

$$\quad = 0 \text{ otherwise}$$

Analytical method

If the variation in time of the pointwise reliability, $Q(t)$, is desired, the above method cannot be used; rather, the distribution of $B_i(0)$ and $B_i^{(d)}(t)$ are approximated by continuous distributions which are assigned from experience. The distribution parameters for the assigned distributions are assumed to be known functions of time. For instance, both $B_i(0)$ and $B_i^{(d)}(t)$ might be assigned normal distributions; the mean and variance of $B_i(0)$ would be constant while those of $B_i^{(d)}(t)$ could be monotonically increasing functions of time. Equation (14) is again used to find the distribution of $B_i(t)$, as shown by the density function in Eq. (20).

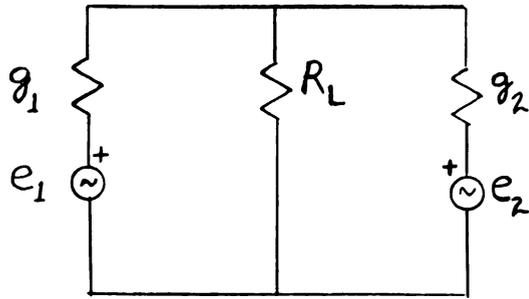
$$(20) \quad f_{B_i}(b_i; t) = \int_{-\infty}^{\infty} f_{B_i}^{(d)}(b_i - z; t) f_{B_i}(z; 0) dz$$

Example

As an example, the pointwise reliability, $Q(t)$, is computed for the d-c power supply system shown in Fig. 14.

The following assumptions are made.

- i) Neither voltage supply can drift; the drift of $B_1(t)$ and $B_2(t)$ is independent.
- ii) Failure of any part parameter removes the corresponding supply from the system; one failure mode is allowed.
- iii) The system is loaded.



State table

X_4	X_3	X_2	X_1	X
0	0	0	0	0
0	0	0	1	1
0	0	1	0	2
0	0	1	1	3
0	1	0	0	4
0	1	1	0	3
1	0	0	0	8
1	0	0	1	3
1	1	0	0	3

$$B_1(t) = g_1(t)$$

$$B_2(t) = g_2(t)$$

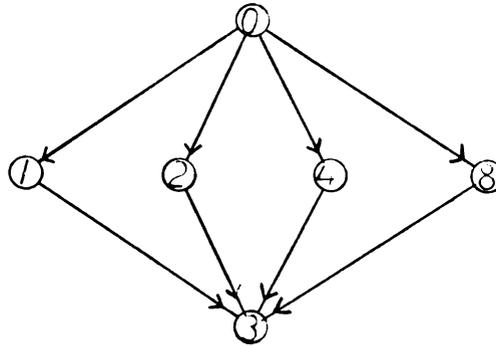
$$B_3(t) = e_1$$

$$B_4(t) = e_2$$

$$R_L = 50 \text{ ohms}$$

$$e_1 = e_2 = e = 100 \text{ volts}$$

Transition diagram



$$Z_0 = (0, 1, 2, 4, 8)$$

$$Z_1 = (3)$$

Fig. 14.--Power supply system and corresponding state table and transition diagram.

iv) The load current, i_L , is the only system parameter.

The system parameter is expressed in terms of the part parameters as follows, using the notation of Eq. (II.3).

$$h_0(\bar{B}(t)) = \frac{e[B_1(t) + B_2(t)]}{1 + R_L[B_1(t) + B_2(t)]}$$

$$h_1(\bar{B}(t)) = h_8(\bar{B}(t)) = \frac{eB_2(t)}{1 + R_L B_2(t)}$$

$$h_2(\bar{B}(t)) = h_4(\bar{B}(t)) = \frac{eB_1(t)}{1 + R_L B_1(t)}$$

System state variable

The distribution of the system state variable for the loaded system is stated below.

$$p_X(0;t) = p_{00}(t) = \exp(-q_0 t)$$

$$p_X(j;t) = p_{0j}(t) = \frac{q_{0j}}{q_j - q_0} \left[\exp(-q_0 t) - \exp(-q_j t) \right]$$

$$j = 1, 2, 4, 8$$

The numerical failure rates are given below in dimensions of (hours)⁻¹, using the notation of Eq. (C.1).

$$\begin{aligned} q_{01} = H_{01} &= 2 \cdot 10^{-5} & H_{21} = H_{81} &= 4 \cdot 10^{-5} \\ q_{02} = H_{02} &= 3 \cdot 10^{-5} & H_{12} = H_{42} &= 2 \cdot 10^{-4} \\ q_{04} = H_{03} &= 10^{-4} & H_{23} = H_{83} &= 3 \cdot 10^{-4} \\ q_{08} = H_{04} &= 5 \cdot 10^{-5} & H_{14} = H_{44} &= 15 \cdot 10^{-5} \end{aligned}$$

$$q_0 = H_{01} + H_{02} + H_{03} + H_{04} = 2 \cdot 10^{-4}$$

$$q_2 = H_{21} + H_{23} = q_8 = H_{81} + H_{83} = 34 \cdot 10^{-5}$$

$$q_1 = H_{12} + H_{14} = q_4 = H_{42} + H_{44} = 35 \cdot 10^{-5}$$

The value of this density function for the states in Z_0 is given below at different times.

<u>t (hours)</u>	<u>$P_X(0;t)$</u>	<u>$p_X(1;t)$</u>	<u>$p_X(2;t)$</u>	<u>$p_X(4;t)$</u>	<u>$p_X(8;t)$</u>
0	1	0	0	0	0
100	0.9802	0.00195	0.00126	0.00970	0.00210
500	0.9048	0.00863	0.0131	0.0432	0.02185
1000	0.8187	0.0152	0.0229	0.0760	0.0382
5000	0.3679	0.0265	0.0397	0.1327	0.0662

Drift analysis

The drift properties of the two conductances are assumed to be the same and are given below, using the notation of Eq. (16).

Both $B_i(0)$ and $B_i^{(d)}(t)$, $i=1,2$, have normal distributions, the former with mean m_0 and standard deviation s_0 and the latter with mean $m_d(t)$ and standard deviation $s_d(t)$. The following values are in mhos.

$$\begin{aligned} m_0 &= 1 \quad , \quad s_0 = 0.0333 \\ m_d(t) &= 0.5 \left[1 - \exp(-8 \cdot 10^{-3} t) \right] \\ s_d(t) &= 0.5 \left[\exp(-2 \cdot 10^{-3} t) - 1 \right] \end{aligned}$$

The mean decreases and the standard deviation increases in time. For all values of time, the 3s limits are positive so that negative values of conductance are, for all practical purposes, eliminated.

The tolerance set is taken as $S(t) = (1.95, \infty)$. To compute $Q(t)$, the integrals in Eq. (1c) are expressed in terms of cumulative distribution functions.

$$\int_{1.95}^{\infty} f_{A/X}(a/x;t) da = 1 - F_{A/X}(1.95/x;t)$$

In state 0, $A(t) = h_0(t)$. The sum $B_1(t) + B_2(t)$ has a normal distribution with mean $m(t)$ and standard deviation $s(t)$.

$$\begin{aligned} m(t) &= 2 \left[m_0 + m_d(t) \right] \\ s(t) &= \left[2s_0^2 + 2s_d^2(t) \right]^{1/2} \end{aligned}$$

The cumulative distribution function of A at any time t can thus be expressed as follows.

$$1 - r_{A/X}(1.95/0;t) = 1 - P \left[\frac{B_1(t) + B_2(t) - m(t)}{s(t)} \leq \frac{0.78 - m(t)}{s(t)} \right]$$

$$= \Phi \left[\frac{m(t) - 0.78}{s(t)} \right] \equiv W_0(t)$$

In the above expression, Φ is the cumulative distribution function of the standardized normal distribution for which tables are available.

Since the drift properties of the conductances are identical, the values of the integrals in Eq. (1c) are the same for states 1,2,4,8.

$$1 - F_{A/X}(1.95/j;t) = \Phi \left[\frac{\frac{1}{2}m(t) - 0.78}{(2)^{1/2}s(t)} \right] \equiv W_1(t), \quad j=1,2,4,8$$

Numerical values for $W_0(t)$ and $W_1(t)$ are tabulated below.

<u>t</u>	<u>$W_0(t)$</u>	<u>$W_1(t)$</u>
0	1.0000	1.0000
100	1.0000	0.9986
500	1.0000	0.0281
1000	1.0000	0.0002
5000	0.9999	0.0000

Finally, $Q(t)$ is computed from Eq. (1c) as shown below.

$$Q(t) = W_0(t) p_X(0;t) + W_1(t) \left[p_X(1;t) + p_X(2;t) + p_X(4;t) + p_X(8;t) \right]$$

<u>t</u>	<u>$Q(t)$</u>
0	1
100	0.9952
500	0.9074
1000	0.8218
5000	0.3679

It should be noted that if the catastrophic failure effects were neglected, the operational effectiveness of the system would be judged by $W_0(t)$; if drift were neglected, the

reliability, $R(0,t)$, is the sum of the density function values of the system state variable at t for all non-failed states. In particular,

$$R(0,5,000) = 0.6330.$$

Both these procedures give optimistic predictions, when compared to $Q(t)$.

Dependent drift

The second type of drift to be discussed is called dependent drift. In this case, the drift characteristics of the part parameters depend not only on which states the system assumes but also on the length of time the system remains in each state.

For any $x \in Z_0$ and any t , Eq. (15) is again assumed to be satisfied so that the drift properties of each part parameter can be investigated separately. To find the distribution of an operational part parameter, given that the system is in state $x \in Z_0$, each possible sequence of states through which the system can pass, starting at state 0 and ending at state x , must be investigated. This is accomplished by defining an auxiliary random variable, called the sequence variable and denoted by Y_x , for each state of the system in Z_0 .

A system is said to assume sequence $(\bar{x}, \bar{t}; x, t)$, where $\bar{x} = (x_1, \dots, x_N)$ and $\bar{t} = (t_0, t_1, \dots, t_N)$, if the state-to-state transitions, beginning in state 0 and ending in state x , occur in the following order. The system goes from state 0 to state x_1 at time t_0 , from x_1 to x_2 at t_1 , from x_{N-1} to x_N at t_{N-1} , from x_N to x at t_N , and the system remains in state x from

t_N to t , where $0 < t_1 < t_2 < \dots < t_N < t$. The sequence variable is assigned a unique value for each sequence the system can assume.

Assuming that $B_i(t)$ is operational when the system is in state x , its distribution is expressed in terms of the sequence variable for state x by Eq. (21).

$$(21) \quad f_{B_i/X}(b_i/x;t) = \sum_y \int_{\bar{t}} f_{B_i/Y_x}(b_i/y;t) f_{Y_x}(y;t) d\bar{t}$$

Each term in Eq. (21) represents one sequence which the system can assume in going from state 0 to state x Z_0 . The value y of Y_x specifies the exact transition times for the sequence. Integration is over all possible transition times, $0 < t_0 < \dots < t_N < t$.

The conditional density term $f_{B_i/X}$ is found from an extension of the independent drift development. The amount by which part parameter $B_i(t)$ drifts when the system is in state j for a time interval of length T is denoted by the random variable $B_{ji}^{(d)}(T)$. This is the amount of change in $B_i(t)$ from its value when the system enters state j . If the system assumes the sequence $((x_1, \dots, x_N), (t_0, \dots, t_N); x, t)$, the relationship between the value of $B_i(t)$ at time t and the drift variables is given by Eq. (22).

$$(22) \quad B_i(t) = B_i(0) + B_{0i}^{(d)}(t_0) + B_{x_1i}^{(d)}(t_1 - t_0) + \dots \\ + B_{x_{N-1}i}^{(d)}(t_N - t_{N-1}) + B_{xi}^{(d)}(t - t_N)$$

The drift in each state is assumed to be independent of the value of $B_i(t)$ when the system enters that state. Thus, the

random variables on the right of Eq. (22) are statistically independent for any set of times, $0 < t_0 < \dots < t_N < t$. If the system assumes the sequence used in Eq. (22), the distribution of $B_i(t)$ is found from Eq. (14) as shown by the density function in Eq. (23). For convenience, the density function of $B_{ji}^{(d)}(T)$ is written as $f_{ji}(z;t)$ and that of $B_i(0)$, as $f_i(z;0)$.

$$(23) \quad f_{B_i/Y_x}(b_i/y, \bar{t}; t) = \int_{\bar{z}} f_{xi} \left[b_i - \sum_{j=0}^{N+1} z_j; t-t_N \right] f^{(d)}(\bar{x}; \bar{t}) d\bar{z}$$

The following notation is used in Eq. (23).

$$f^{(d)}(\bar{x}; \bar{t}) = f_{x_{N+1}}(z_{N+1}; t_N - t_{N-1}) \dots f_{0i}(z_1; t_0) f_i(z_0; 0)$$

$$d\bar{z} = dz_{N+1} dz_N \dots dz_1 dz_0$$

The integration is over all values of the z 's for which the corresponding density functions are non-zero.

Since the drift variables in Eq. (23) are independent, the central limit theorem can be used to approximate the distribution of $B_i(t)$ by a normal distribution. If the distributions of the variables in Eq. (22) are near-normal, this would be a good approximation. However, for arbitrary distributions, the number of variables must be large before this approximation is valid.

The distribution of the sequence variable for a particular state is now discussed. The lifetime variables for all part parameters are assumed to have exponential distributions which are given in Eq. (III.9), for non-loaded systems, and in Eq. (C.1), for loaded systems. To minimize the notational com-

plexity, each part parameter is assumed to have only one failure mode. If the system assumes the sequence $((x_1, \dots, x_N), (t_0, \dots, t_N); x, t)$, the part parameter that fails at time t_j is denoted by $B_{s_j}(t)$, $j=0,1,\dots,N$.

If the system is non-loaded, $f_{Y_x}(y;t)d\bar{t}$ is computed as follows.

i) The probability that $B_{s_j}(t)$ fails at time t_j , $j=0,1,\dots,N$, is the following.

$$\prod_{j=0}^N H_{s_j} \exp \left[- \sum_{j=0}^N H_{s_j} t_j \right] d\bar{t}$$

ii) The probability that no part parameter in $\bar{B}_x(t)$ fails to time t is the following.

$$\exp \left[-t \sum_{i=1}^x H_{r_j} \right]$$

Both events must be realized so that $f_{Y_x}(y;t)dt$ is the product of the two terms.

If the system is loaded, the failure rates become functions of the states and the expression for $f_{Y_x}(y;t)d\bar{t}$ becomes more complex, as shown below.

i) The probability that $B_{s_j}(t)$ fails at time t_j is the following, where x_0 represents state 0.

$$\prod_{j=0}^N H_{x_j s_j} \exp - \left[\sum_{j=0}^N H_{x_j s_j} t_j \right]$$

ii) The probability that no part parameter in $\bar{B}_x(t)$ fails

to time t_0 is the following.

$$\exp\left[-t_1 \sum_{j=1}^x H_{Or_j}\right]$$

iii) The probability that no part parameter in $\bar{B}_x(t)$ fails from time t_j to t_{j+1} , $j=0,1,\dots,N-1$ is the following.

$$\exp\left[-(t_{j+1}-t_j) \sum_{i=1}^x H_{x_j r_i}\right]$$

iv) The probability that no part parameter in $\bar{B}_x(t)$ fails from time t_N to t is the following.

$$\exp\left[-(t-t_N) \sum_{i=1}^x H_{x_N r_i}\right]$$

The product of all terms above is the required density function.

In stating the expressions above, it was assumed that no part parameter is removed from a system because of the failure of other part parameters. If this assumption is not satisfied, extra factors must be inserted into some of the terms in Step iii.

Conclusions

The failure analysis used to analytically predict the susceptibility of a redundant system to failure is dependent on three factors, viz., the failure state, the figures of merit, and the available time and money. The General Model furnishes a practical means for accomplishing a failure analysis at the level of sophistication and complexity that is dictated by these factors.

If only catastrophic failure data is available, the one-failure and two-failure algorithms permit direct computation of reliability when the system is non-loaded. If drift data is also available, the General Model partitions the failure analysis into a set of drift problems which are re-united on the basis of catastrophic failure tendencies. A higher level of complexity is introduced when the system is loaded and when the drift is dependent. Correspondingly, both the amount of required data and the analytical complexity increase.

Although only non-repairable systems are considered in this thesis, the theory of Markov chains used with loaded systems is directly adaptable to repairable systems. The difficulty of the drift problem is magnified in repairable systems, especially if the drift is dependent.

Two extensions of the work in this thesis would be of great value. Firstly, the distribution of the system state variable for loaded systems should be studied when the lifetimes do not have exponential distributions. Secondly, methods for finding the joint distribution of operational part parameters

from their drift properties are needed when they are not statistically independent.

APPENDIX A

SOME BASIC DEFINITIONS FROM PROBABILITY THEORY

These definitions are taken from Doob (56).

Definition: Probability set function P

Let Ω be some basic space; let there be a certain collection of sets of points of Ω which are called measurable sets. The class of measurable sets is assumed to be a Borel field. A (set) function P is assumed to be defined for all measurable sets and is a probability measure, i.e., P is completely additive, non-negative, and $P(\Omega) = 1$.

The number $P(A)$ is called the probability or the measure of the (measurable) set A of points of Ω , or w points.

Definition: Random variable

A (real) function x , defined on a space of w points, is called a (real) random variable if there exists a probability measure P defined on the w sets and if, for every real number a , the inequality $x(w) \leq a$ delineates a measurable w set; i.e., $F(A)$ is defined for all real a .

$$F(a) = P[x(w) \leq a]$$

Thus, a (real) random variable is a (real) measurable function.

Definition: Stochastic process

A stochastic process is any family of random variables

$[x(t), t \in T]$. The notation used is that $x(t)$ is the observation at time t and T is the time range involved.

The term "stochastic process" is usually applied only when the process involves infinitely many random variables. Historically, this term has been reserved for families of random variables with some simple relationship between the variables.

Definition: Sample function of a stochastic process

If $[x(t), t \in T]$ is a stochastic process, a function of $t \in T$ obtained by fixing w in $x(t,w)$ and letting t vary is called a sample function of the process, where $x(t,w)$ is the value of the random variable $x(t)$ at the point w .

Definition: Markov process and Markov chain

A (strict sense) Markov process is a stochastic process $[x(t), t \in T]$ which satisfied the following condition.

For any integer $n \geq 1$, if $t_1 < t_2 < \dots < t_n$ are parameter values, the conditional probability distribution of $x(t_n)$ relative to $x(t_1), \dots, x(t_{n-1})$ is the same as that relative to $x(t_{n-1})$ in the sense that for each (real) a , the following is satisfied with probability 1.

$$P[X(t_n) \leq a / x(t_1), \dots, x(t_{n-1})] = P[x(t_n,) \leq a / x(t_{n-1})]$$

Whenever Markov processes are discussed, strict sense is assumed. If the values of $x(t)$ for any t are discrete in nature, the Markov process is called a Markov chain.

Definition: Separable stochastic processes

Let $[x(t), t \in T]$ be a real stochastic process with linear parameter set T . Let Ψ be a system of linear Borel sets. Then,

the process is called separable relative to Ψ if the following condition is true.

Let $\{t_j\}$ be a sequence of parameter values; let A be a w set of probability zero; let I be any open interval.

Then, the w sets, $[w: x(t,w) \in B, t \in IT]$,
 $[w: x(t_j,w) \in B, t_j \in IT]$, where $B \in \Psi$, differ by at most a subset of A .

If Ψ is the class of all (finite or infinite) closed intervals, the process is called separable.

If T is an infinite sequence, the stochastic process $[x(t), t \in T]$ is called a discrete random process or a random sequence. The general sample is a sequence. If T is an interval, the stochastic process is called a continuous parameter process. The general sample is a function of t defined on an interval.

Markov chains which are random sequences are studied in detail by Feller (57). Both random sequences and continuous parameter Markov processes are considered by Bartlett (58). The main interest here is the continuous parameter process, considered in Appendix B.

APPENDIX B

CONTINUOUS PARAMETER MARKOV CHAINS WITH A FINITE NUMBER OF STATES

This discussion is taken mainly from Doob (56).

Let $[x(t), 0 \leq t \leq \infty]$ be a (continuous parameter) Markov chain such that the random variables $x(t)$ assume the values $1, 2, \dots, N$. Each value is called a state of the chain.

If $P(x(s, \omega) = i) > 0$, the probability, given state i at time s , that state j is found at time t is called a transition probability and written as $p_{ij}(s, t)$.

$$p_{ij}(s, t) = P(x(t, \omega) = j / x(s, \omega) = i) , t \geq s$$

Two fundamental properties of transition probabilities are stated in Eq. (B.1) and Eq. (B.2).

Let $P(x(s, \omega) = i) > 0$.

$$(B.1) \quad p_{ij}(s, t) \geq 0, \quad \sum_j p_{ij}(s, t) = 1$$

$$(B.2) \quad p_{ik}(s, u) = \sum_j p_{ij}(s, t) p_{jk}(t, u), \quad 0 \leq s < t < u$$

Equation (B.2) is a special case of the Chapman-Kolmogorov equation. The sum is over those values of j for which $p_{jk}(t, u)$ is defined.

Definition: Markov transition matrix function

Let $\bar{P}(s, t)$ be the matrix with typical element $p_{ij}(s, t)$.

If $\bar{P}(s, t)$ satisfied Eq. (B.1) and Eq. (B.2) it is called a Markov transition matrix function.

A sufficient condition for the existence of a Markov process is now stated.

Given a Markov transition matrix function, there exists a corresponding Markov process $[x(t), 0 \leq t < \infty]$ which is obtained as follows.

i) Choose any initial probability distributions.

$$P(x(0, \omega) = i) , i = 1, 2, \dots, N$$

ii) For every finite t set, $0 = t_0 < t_1 < \dots < t_n$,

define the following.

$$\begin{aligned} &P(x(t_0, \omega) = a_0, x(t_1, \omega) = a_1, \dots, x(t_n, \omega) = a_n) \\ &= P(x(0, \omega) = a_0) P(x(t_1, \omega) = a_1 / x(0, \omega) = a_0) \dots \\ &\quad P(x(t_n, \omega) = a_n / x(t_{n-1}, \omega) = a_{n-1}) \\ &= P(x(0, \omega) = a_0) p_{a_0 a_1}(0, t_1) \dots p_{a_{n-1} a_n}(t_{n-1}, t_n) \end{aligned}$$

If, for each pair (i, j) the transition probability $p_{ij}(s, t)$ depends only on $t-s$ whenever $P(x(s, \omega) = i) > 0$, then the Markov process is said to have stationary transition probabilities.

The two fundamental properties given in Eq. (B.1) and Eq. (B.2) then become those in Eq. (B.3) and Eq. (B.4).

$$(B.3) \quad p_{ij}(t) \geq 0, \quad \sum_j p_{ij}(t) = 1, \quad t > 0$$

$$(B.4) \quad p_{ik}(s+t) = \sum_j p_{ij}(s) p_{jk}(t), \quad s, t > 0$$

Definition: Stationary Markov transition matrix function

Let $\bar{P}(t)$ be the matrix with typical element $p_{ij}(t)$. If $\bar{P}(t)$ satisfied Eq. (B.3) and Eq. (B.4), it is called a stationary Markov transition matrix function.

A Markov chain defined in terms of a stationary Markov transition matrix function, with some initial probability distribution, will be called a stationary Markov chain. Only this type of chain will be discussed.

Solutions for the transition probabilities are discussed below. Two general approaches may be used. The first considers equations resulting from the differentiation of Eq. (B.4). The resulting solutions are usually presented in matrix form. The second method, summarized below, attacks the problem from the viewpoint of the sample function. It is felt that this second method leads to greater understanding because of its more direct probabilistic interpretation.

Definition: Step function

A function $g(t)$ will be called a step function if it satisfies all of the following conditions.

i) $g(t)$ has only finitely many points of discontinuity in every finite closed interval;

ii) $g(t)$ is identically constant in every open interval of continuity points;

iii) If t_0 is a point of discontinuity, either of the inequalities in Eq. (B.5) is satisfied.

$$(B.5) \quad g(t_0^-) \leq g(t_0) \leq g(t_0^+)$$

$$g(t_0^-) \geq g(t_0) \geq g(t_0^+)$$

Definition: Jump of a function

A function $g(t)$ is said to have a jump at the point t_0 if it is discontinuous there, and the one sided limits $g(t_0^-)$ and $g(t_0^+)$ exist and satisfy one of the inequalities in Eq. (B.5).

Before considering solutions for the transition probabilities, some useful properties concerning stationary Markov chains are given. The following assumptions are used.

If $\bar{P}(t) = [p_{ij}(t)]$ is a stationary Markov transition matrix function, it is assumed that the conditions in Eq. (B.6) is satisfied, where $p_{ij}(t)$ is continuous when $t = 0$.

$$(B.6) \quad \lim_{t \rightarrow 0} p_{ij}(t) = 1 \quad \text{for } i = j$$

The following conclusions are evident from Eq. (B.6).

- i) $\lim_{t \rightarrow 0} p_{ij}(t) = 0$ for $i \neq j$.
- ii) $p_{ij}(t)$ is continuous for all t .
- iii) If $j \neq i$, $p_{ij}(t)$ either vanishes identically or never vanishes, except when $t = 0$.

Theorem B.1

Let $[p_{ij}(t)]$ be a stationary Markov transition matrix function. Then

$\lim_{t \rightarrow \infty} p_{ij}(t)$ exists for all i, j and the limit is approached exponentially fast.

Theorem B.2

Let $[p_{ij}(t)]$ be a stationary Markov transition matrix function satisfying Eq. (B.6). Then, the limit in Eq. (B.7) exists for all i .

$$(B.7) \quad \lim_{t \rightarrow 0} \frac{1 - p_{ii}(t)}{t} = q_i$$

Furthermore, if $[x(t), 0 \leq t < \infty]$ is a separable process determined by $[p_{ij}(t)]$, together with an initial distribution, Eq. (B.8) is satisfied.

$$(B.8) \quad p \left[x(\tau, w) = i, t_0 \leq \tau \leq t_0 + a \mid x(t_0, w) = i \right] = \exp(-q_i a)$$

Theorem B.3

Let $[p_{ij}(t)]$ be a stationary Markov transition matrix function satisfying Eq. (B.6). Then, the limits in Eq. (B.9) exist and satisfy Eq. (B.10).

$$(B.9) \quad \lim_{t \rightarrow 0} \frac{p_{ij}(t)}{t} = q_{ij}, \quad i \neq j,$$

$$(B.10) \quad \sum_{j \neq i} q_{ij} = q_i \geq 0$$

If, in addition, $[x(t), 0 \leq t < \infty]$ is a separable process determined by $[p_{ij}(t)]$ together with an initial distribution, then there exists with probability 1 a sample function discontinuity which is a jump. In fact, there exists a first and a last discontinuity which are jumps.

Theorem B.4

The sample function of a separable Markov chain with a finite number of states having stationary transition probabilities satisfying Eq. (B.6) are almost all step functions.

Equations (B.7), (B.8), and (B.9) express three crucial points. The probability of no transition is seen to be asymptotically linear as t approaches 0 from Eq. (B.7). From Eq. (B.8), the probability of no transition in some interval is seen to be a decreasing exponential function of the length of the interval. Finally, Eq. (B.9) shows that the probability of transition from one state to any other is also asymptotically linear as t approaches 0. These facts are used to derive the equations describing the transition probabilities.

Let ${}_1P_{ik}(t)$ be the probability, given $x(t_0, w) = i$, that $X(t_0+t, w) = k$ and the transition from i to k is accomplished in exactly one step. This implies, from Theorem B.4, that the sample function is a step function which takes on the two values i and k and that there is only one jump discontinuity in the sample function. This jump point is labelled s .

Then, ${}_1P_{ik}(t)$ can be evaluated by the following considerations.

i) $x(\tau, w) \equiv i$ to the jump point s . From Eq. (B.8), the probability of this event is $\exp(-q_1 s)$.

ii) There is a jump to k at the jump point s . From Eq. (B.9), the probability of this is $q_{ik} ds$.

iii) $x(\tau, w) \equiv k$ from s to t . From Eq. (B.8), this probability is $\exp(-q_1 s)$.

iv) The probability of realizing all three of these events, for any s in the interval $(0, t)$ is given below.

$${}_1P_{ik}(t) = \int_0^t q_{ik} \exp[-q_1 s - q_k(t-s)] ds \quad k \neq i$$

$$= 0 \quad k = i$$

Now let ${}_n P_{ik}(t)$ be the probability, given $x(t_0, w) = i$, that $x(t_0+t, w) = k$ and that the transition from i to k has been accomplished in exactly n steps. An induction process is used to derive this probability. However, two methods can be used to state the induction process, depending on how the sample function is viewed. In both methods, Eq. (B.11), which is based on Eq. (B.8), is used

$$(B.11) \quad {}_0 P_{ik}(t) = 0 \quad , \quad k \neq i$$

$$= \exp(-q_1 t) \quad , \quad k = i$$

The first jump method

Here, the sample function is viewed as having a first jump at s . The following considerations lead to the general term of the induction process.

i) The probability of no jump until s is $\exp(-q_i s)$ from Eq. (B.8).

ii) The probability of a jump at s from state i to any state $j \neq i$ is $q_{ij} ds$ from Eq. (B.9).

iii) The probability of going from j to the final state of interest k in exactly n steps is ${}_n p_{jk}(t-s)$.

iv) The above three events can be realized for any state $j \neq i$. Thus, the general term of the induction process is given by Eq. (B.12).

$$(B.12) \quad {}_{n+1} p_{ik}(t) = \sum_{j \neq i} \int_0^t q_{ij} \exp(-q_i s) {}_n p_{jk}(t-s) ds, \quad n \geq 0$$

The last jump method

In this method, the last jump of the sample function is assumed to take place at s . The general term of the induction process is derived as shown below.

i) The probability of going from the initial state i to any other state $j \neq k$ in exactly n steps is ${}_n p_{ij}(s)$.

ii) The probability of jumping to state k from state $j \neq k$ at the last jump s is $q_{jk} ds$.

iii) The probability of no jump between s and t is $\exp[-q_k(t-s)]$.

iv) The above events can be realized for any state $j \neq k$.

Thus, the general term of the induction process is given by Eq. (B.13).

$$(B.13) \quad {}_{n+1}p_{ik}(t) = \sum_{j \neq k} \int_0^t q_{jk} \exp -q_k(t-s) {}_n p_{ij}(s) ds, \quad n \geq 0$$

Since, by Theorem (B.4), almost all sample functions are step functions, a general expression for any stationary transition probability is given by Eq. (B.14).

$$(B.14) \quad p_{ik}(t) = \sum_{n=0}^{\infty} {}_n p_{ik}(t)$$

The equations for either of the induction processes together with Eq. (B.14) provide an explicit algorithm by which, given the q_i 's and q_{ij} 's, the transition probabilities can be computed. Specifically, the two algorithms are represented by Eq.'s (B.15) and (B.16).

$$(B.15) \quad p_{ik}(t) = \delta_{ik} \exp(-q_i t) + \sum_{j \neq i} \int_0^t q_{ij} \exp(-q_i s) p_{jk}(t-s) ds$$

Equation (B.16) is obtained from Eq.'s (B.11), (B.13), and (B.14).

$$(B.16) \quad p_{ik}(t) = \delta_{ik} \exp(-q_i t) + \sum_{j \neq k} \int_0^t q_{jk} \exp[-q_k(t-s)] p_{ij}(s) ds$$

In Eq. (B.15) and Eq. (B.16), δ_{ik} is the Kronecker delta function defined below.

$$\begin{aligned} \delta_{ik} &= 1, \quad i = k \\ &= 0, \quad i \neq k \end{aligned}$$

An alternate method of solution is found from the derivatives of Eq. (B.15) and Eq. (B.16).

The set of equations in Eq. (B.17), called the "backward system", results from differentiating Eq. (B.15).

$$(B.17) \quad p_{ik}'(t) = -q_i p_{ik}(t) + \sum_{j \neq i} q_{ij} p_{jk}(t) ; i, k = 1, \dots, N$$

The set of equations in Eq. (B.18), called the "forward system", results from differentiating Eq. (B.16).

$$(B.18) \quad p_{ik}'(t) = -p_{ik}(t) q_k + \sum_{j \neq k} q_{jk} p_{ij}(t) ; i, k = 1, \dots, N$$

The initial conditions for both systems are shown in Eq. (B.19).

$$(B.19) \quad \begin{aligned} p_{ij}(0) &= 1 & , & \quad i = j \\ &= 0 & \quad i \neq j \end{aligned}$$

From Eq. (B.17) or Eq. (B.18) together with Eq. (B.19), it can be seen that the q 's and q_{ij} 's determine the transition probabilities uniquely. It can be shown that Eq. (B.17) and Eq. (B.18) with initial conditions in Eq. (B.19) have a unique solution satisfying Eq. (B.1) and Eq. (B.2) if Eq. (B.10) is satisfied. The q_i 's and q_{ij} 's themselves must be known. However, they should be deducible either experimentally or theoretically using Eq. (B.7) and Eq. (B.9). That is, from Eq. (B.7), the probability of no transitions in time dt is $1 - q_i dt$, where second order terms are neglected. Similarly, the probability of a transition from i to j in time dt is $q_{ij} dt$, again neglecting second order terms.

APPENDIX C

THE APPLICATION OF MARKOV CHAINS TO THE ANALYSIS OF LOADED SYSTEMS

In this appendix, the q 's which describe a Markov chain are expressed in terms of the lifetime distributions of part parameters for loaded systems. The part parameters are assumed to have one failure mode. The lifetime density function f_{ir} for part parameter $B_r(t)$ when the system is in state i is assumed to be of the form shown in Eq. (C.1).

$$(C.1) \quad f_{ir}(z) = H_{ir} \exp(-H_{ir} z) \quad \text{if } z \geq 0 \text{ and } i \in Z_0 \text{ and } X_r(t) = 0 \\ = 0 \quad \text{if } z < 0 \text{ or } i = Z_1 \text{ or } X_r(t) = 1$$

The q_i constants defined in Eq. (B.7) are considered first.

$$q_i = \lim_{t \rightarrow 0} \frac{1 - p_{ii}(t)}{t}$$

If $i = Z_1$, $p_{ii}(t) \equiv 1$, since the system cannot leave the failed state, so that $q_i = 0$.

If $i \in Z_0$, and if $X_{r_m}(t) = 0$, $m = 1, 2, \dots, n$ when the system is in state i , the system stays in state i if none of the part parameters B_{r_m} fails in the time interval of length t . Then, $p_{ii}(t)$ can be expressed as follows.

$$\begin{aligned}
 p_{ii}(t) &= \prod_{m=1}^n \left[\int_t^{\infty} f_{ir_m}(z) dz \right] \\
 &= \prod_{m=1}^n \left[\exp(-H_{ir_m} t) \right] = \exp(-H_e t)
 \end{aligned}$$

$$H_e = \sum_{m=1}^n H_{ir_m}$$

The expression for q_i is then obtained by substituting $p_{ii}(t)$ into the equation defining q_i .

$$q_i = \lim_{t \rightarrow 0} \frac{1 - \exp(-H_e t)}{t} = \lim_{t \rightarrow 0} \left[H_e \exp(-H_e t) \right] = H_e$$

Thus, q_i is the sum of the failure rates of all part parameters which are not failed when the system is in state i . The expressions for q_i are summarized in Eq. (C.2).

$$\begin{aligned}
 \text{(C.2)} \quad q_i &= \sum_{m=1}^n H_{ir_m} && \text{if } i \in Z_0 \\
 &= 0 && \text{if } i = Z_1
 \end{aligned}$$

The expression for q_{ij} , $i \neq j$ defined in Eq. (B.9) is now considered.

$$q_{ij} = \lim_{t \rightarrow 0} \frac{p_{ij}(t)}{t}, \quad i \neq j$$

If $i, j \in Z_0$, and if j can be reached from i in one step, then there is one part parameter, say $B_{r_1}(t)$, for which $X_{r_1}(t) = 0$ in i and $X_{r_1}(t) = 1$ in j . It is assumed that $X_{r_m}(t) = 0$, $m = 2, 3, \dots, n$, in both i and j . The expression for q_{ij} is obtained by means of Eq. (B.14).

$$p_{ij}(t) = \sum_{n=0}^{\infty} n p_{ij}^{(n)}(t)$$

Since $i \neq j$, ${}_0 p_{ij}(t) = 0$.

Since the system has only a finite number of states and since a state cannot be reached after the system has left it, this sum has only a finite number of terms, say N . Then, q_{ij} can be expressed as follows.

$$q_{ij} = \lim_{t \rightarrow 0} (1/t) \sum_{n=1}^N n p_{ij}^{(n)}(t) = \sum_{n=1}^N \lim_{t \rightarrow 0} [n p_{ij}^{(n)}(t)/t]$$

The term ${}_1 p_{ij}(t)$ is computed below. The dummy variable s is some time for which the system is in state i , but which is less than t .

i) The probability that the part parameter $B_{r_1}(t)$ fails at time s is the following.

$$f_{ir_1}(s) ds = H_{ir_1} \exp(-H_{ir_1} s) ds$$

ii) The probability that no other part parameter fails to time s is the following.

$$\prod_{m=2}^n \int_s^{\infty} f_{ir_m}(z) dz = \exp(-H_1 s)$$

$$H_1 = \sum_{m=2}^n H_{ir_m}$$

iii) The probability that no other part parameter fails from s to t is the following.

$$\prod_{m=2}^n \left[\int_{t-s}^{\infty} f_{jr_m}(z) dz \right] = \exp[-H_2(t-s)]$$

$$H_2 = \sum_{m=2}^n H_{jr_m}$$

These three events can be realized for any s .

$${}_1P_{ij}(t) = \int_0^t H_{ir_1} \exp(-H_e s) \exp[-H_2(t-s)] ds$$

$$H_e = H_1 + H_{ir_1} = \sum_{m=1}^n H_{ir_m} \neq H_2$$

$${}_1P_{ij}(t) = \left[H_{ir_1} / (H_2 - H_e) \right] \left[\exp(-H_e t) - \exp(-H_2 t) \right]$$

The first term in the expression for q_{ij} can now be found.

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{{}_1P_{ij}(t)}{t} &= \lim_{t \rightarrow 0} \left(\left[H_{ir_1} / (H_2 - H_e) \right] \left[-H_e \exp(-H_e t) + H_2 \exp(-H_2 t) \right] \right) \\ &= H_{ir_1} \end{aligned}$$

The following can be shown.

$$\lim_{t \rightarrow 0} \frac{{}_n P_{ij}(t)}{t} = 0 \text{ for } n \geq 2$$

Thus, q_{ij} is given by Eq. (C.3).

$$(C.3) \quad q_{ij} = H_{ir_1}$$

That is, if non-failed state j can be reached from state i in one step, q_{ij} is the failure rate of the part parameter whose failure would cause the transition from i to j .

If $j = Z_1$ and if j can be reached from i in one step, there may be a set of part parameters such that if any one fails,

the system goes from i to Z_1 . If this set of part parameters is B_{r_1}, \dots, B_{r_v} , q_{ij} is given by Eq. (C.4).

$$(C.4) \quad q_{ij} = \sum_{m=1}^v H_{ir_m}$$

APPENDIX D

JUSTIFICATION OF THE TWO-FAILURE ALGORITHM

When all part parameters can fail in either of two modes, and when drift is neglected, the two-failure algorithm presented in Chapter III can be used to find reliability expressions. The theoretical justification for this algorithm is demonstrated in this appendix.

The four terms used in the two-failure algorithm are defined below. The system is assumed to have k part parameters.

$$p^{(1)}(1;t) = P(A(t) = 1 / X_j(t) \neq 2, j=1, \dots, k)$$

$$p^{(1)}(0;t) = P(A(t) = 0 / X_j(t) \neq 2, j=1, \dots, k)$$

$$p^{(2)}(1;t) = P(A(t) = 1 / X_j(t) \neq 1, j=1, \dots, k)$$

$$p^{(2)}(0;t) = P(A(t) = 0 / X_j(t) \neq 1, j=1, \dots, k)$$

Since $A(t)$ can assume only the values 0 and 1, Eq. (D.1)

is obtained.

$$(D.1a) \quad p^{(1)}(1;t) = 1 - p^{(1)}(0;t)$$

$$(D.1b) \quad p^{(2)}(1;t) = 1 - p^{(2)}(0;t)$$

Equation (D.2) is obtained from Theorem D.1, given below.

$$(D.2) \quad p_A(1;t) = p^{(1)}(1;t) + p^{(2)}(1;t)$$

Equation (D.2) and the definition of a density function produce the following expression.

$$p_A(0;t) = 1 - p_A(1;t) = 1 - [p^{(1)}(1;t) + p^{(2)}(1;t)]$$

The following is found by substituting Eq. (D.1) into this

relation.

$$p_A(0;t) = p^{(1)}(0;t) + p^{(2)}(0;t) - 1$$

The last two equations are used in the final step of the two-failure algorithm. Thus, the proof of Theorem D.1 completes the justification of this algorithm.

Theorem D.1

Let the value of the density function of $A(t)$ at $A(t) = 1$ be given by Eq. (III.17b).

$$(III.17b) \quad p_A(1;t) = \sum_{Z_1} \prod_{i=1}^k p_{X_i}(x_i;t)$$

The sum in Eq. (III.17b) extends over all possible sets of parameter state variable values for which the system is failed, even if some are sets of measure zero.

Let $p^{(1)}(1;t)$ and $p^{(2)}(1;t)$ be computed as prescribed in Steps 1 and 2 of the two-failure algorithm.

Then, Eq. (D.2) is satisfied.

$$(D.2) \quad p_A(1;t) = p^{(1)}(1;t) + p^{(2)}(1;t)$$

Proof:

The following notation is used in this proof.

A 2-table is a (partial) reliability table which lists all possible sets of parameter state variable values for which $A(t) = 1$ when each part parameter has two failure modes; i.e., each row represents one term in Eq. (III.17b).

A 1-open table is the table used to compute $p^{(1)}(1;t)$ is Step 1 of the two-failure algorithm. Similarly, a 1-short table is used in Step 2 to compute $p^{(2)}(1;t)$.

The following relations are defined.

$$p_A(1;t) = p_A^{(1)}(1;t) + p_A^{(2)}(1;t)$$

$$(D.3a) \quad p_A^{(1)}(1;t) = \sum_1 \prod_{i=1}^k p_{X_i}(x_i;t)$$

$$(D.3b) \quad p_A^{(2)}(1;t) = \sum_2 \prod_{i=1}^k p_{X_i}(x_i;t)$$

The one subscript in Eq. (D.3a) means that the sum extends over all possible sets of parameter state variable values for which the system is failed open; the two subscripts in Eq. (D.3b) cover short system failures. Each row in the 2-table is included in one and only one sum in Eq. (D.3).

The method of proof is to show, on the basis of the reliability tables, that Eq. (D.4) is satisfied.

$$(D.4a) \quad p_A^{(1)}(1;t) = p^{(1)}(1;t)$$

$$(D.4b) \quad p_A^{(2)}(1;t) = p^{(2)}(1;t)$$

Proof for Eq. (D.4a)

For simplicity, the method of proof is demonstrated for the reliability diagram of Fig. (III.5). The table below shows entries in the 1-open and 2-tables which produce corresponding terms on the two sides of Eq. (D.4a). Only that portion of the 2-table for which the system is failed open is shown.

Entry	1-open table (for $p^{(1)}(1;t)$)			2-table (for $p_A^{(1)}(1;t)$)		
	$X_3(t)$	$X_2(t)$	$X_1(t)$	$X_3(t)$	$X_2(t)$	$X_1(t)$
1	1	1	1	1	1	1
2	0	1	1	0 2	1 1	1 1
3	1	0	1	1 1	0 2	1 1
4	1	1	0	1 1	1 1	0 2
5	0	1	0	0 0 2 2	1 1 1 1	0 2 0 2

The following table shows the terms which are computed from the separate entries in the preceding table. For convenience, $p_{X_1}(j;t)$ is written as $p_1(j)$. The terms from the 1-open table are computed by Step 1 of the two-failure algorithm; those from the 2-table, by Eq. (III.17b). The following identity is used.

$$\sum_{j=1}^3 p_{X_1}(j;t) = 1$$

Entry	$p^{(1)}(1;t)$	$p_A^{(1)}(1;t)$
1	$p_3(1)p_2(1)p_1(1)$	$p_3(1)p_2(1)p_1(1)$
2	$(1-p_3(1))p_2(1)p_1(1)$	$(p_3(0)+p_3(2))p_2(1)p_1(1)$ $= (1-p_3(1))p_2(1)p_1(1)$
3	$p_3(1)(1-p_2(1))p_1(1)$	$p_3(1)(p_2(0)+p_2(2))p_1(1)$ $= p_3(1)(1-p_2(1))p_1(1)$
4	$p_3(1)p_2(1)(1-p_1(1))$	$p_3(1)p_2(1)(p_1(0)+p_1(2))$ $= p_3(1)p_2(1)(1-p_1(1))$
5	$(1-p_3(1))(1-p_1(1))$ $\quad \cdot p_2(1)$	$p_3(0)(p_1(0)+p_1(2))p_2(1)$ $+ p_3(2)(p_1(0)+p_1(2))p_2(1)$ $= (1-p_3(0))(1-p_1(1))p_2(1)$

This proof hinges on the fact that if the system is failed open when all the part parameters in a certain set are failed open and the remaining part parameters are non-failed, then the system is failed open no matter what states the remaining part parameters assume. This can be seen by noting that, when the system is failed open, its reliability diagram is separated into two parts which cannot be joined together by assuming that any of the remaining part parameters have failed short. This fact is used to make the correspondences in the first table. For instance, in entry 3, the state of $B_2(t)$ cannot change the fact that the system is failed open. Similarly, in entry 5, the system is failed open for any states of $B_1(t)$ and $B_3(t)$.

A similar proof can be stated for Eq. (D.4b).

APPENDIX E

DERIVATION OF AN APPROXIMATION TO THE MEAN AND VARIANCE OF A
FUNCTION OF RANDOM VARIABLES

In this appendix, an approximation to the variance, s^2 , and mean, m , of the random variable A defined in Eq. (IV.8) is derived for use with the Tchebycheff inequality.

The first step involves the Taylor series expansion of the function $v(b_1, \dots, b_k)$ about a point $(b_1, \dots, b_k) = (z_1, \dots, z_k)$ which is given in Eq. (E.1).

$$(E.1) \quad v(b_1, \dots, b_k) = v(z_1, \dots, z_k)$$

$$+ \sum_{j=1}^{\infty} \left((1/j!) \left[\sum_{r=1}^k (b_r - z_r) \partial / \partial b_r \right]^j v(z_1, \dots, z_k) \right)$$

The notation in Eq. (E.1) is interpreted as follows.

$$\begin{aligned} & \left[\sum_{r=1}^k h_r / b_j \right]^j v(b_1, \dots, b_k) \\ &= \sum' \left(\frac{j!}{c_1! c_2! \dots c_k!} \left[\prod_{i=1}^k (h_i)^{c_i} \right] \frac{v(b_1, \dots, b_k)}{b_1^{c_1} b_2^{c_2} \dots b_k^{c_k}} \right) \end{aligned}$$

The prime on the summation sign indicates that the sum extends over all set of values (c_1, c_2, \dots, c_k) for which the following is satisfied.

$$\sum_{i=1}^k c_i = j \text{ and } 0 \leq c_i \leq j$$

In the second term on the right of Eq. (E.1), the arguments (z_1, \dots, z_k) are used with v to signify that the partial derivatives are evaluated at the point (z_1, \dots, z_k) .

The Taylor series expansion of the function in Eq. (IV.8) about the point $(B_1, \dots, B_k) = (m_1, \dots, m_k)$ is given by Eq. (E.1) if z_i is replaced by m_i , where m_i is the mean of B_i . If all terms for which $j \geq 2$ are neglected, Eq. (E.2) is obtained.

$$(E.2) \quad v(B_1, \dots, B_k) - v(m_1, \dots, m_k) \\ \doteq \sum_{r=1}^k (B_r - m_r) v_r(m_1, \dots, m_k)$$

In Eq. (E.2), the following notation is used.

$$v_r(m_1, \dots, m_k) = (\partial / \partial b_r) v(B_1, \dots, B_k)$$

The partial derivative is evaluated at the point (m_1, \dots, m_k) .

In order to approximate the variance of the random variable A , the number S_A^2 is defined by the expectation in Eq. (E.3).

$$(E.3) \quad S_A^2 = E \left[(v(B_1, \dots, B_k) - v(m_1, \dots, m_k))^2 \right]$$

Equation (E.4) results from substituting Eq. (E.2) into Eq. (E.3) and taking the expectation.

$$(E.4) \quad S_A^2 \doteq \sum_{r=1}^k v_r^2(m_1, \dots, m_k) s_r^2 \\ + 2 \sum'' v_{r_1}(m_1, \dots, m_k) v_{r_2}(m_1, \dots, m_k) \text{Cov}(B_{r_1}, B_{r_2})$$

In Eq. (E.4), s_r^2 is the variance of B_r and the double prime implies that the sum extends over all terms for which the following is satisfied.

$$r_2 > r_1 \text{ and } 1 \leq r_1, r_2 \leq n$$

If the B's are statistically independent, all covariance terms are zero.

In order to use Eq. (E.4) as an estimate of the variance of A, the approximation in Eq. (E.5) must be made.

$$(E.5) \quad v(m_1, \dots, m_k) \doteq E[v(B_1, \dots, B_k)]$$

Equation (E.5) is exactly satisfied if v is a linear function. For some non-linear functions, this might produce an extremely poor approximation, as shown by the following example.

$$A = v(B) = B^2$$

If B has a normal distribution with mean zero, then the left side of Eq. (E.5) is zero for the B^2 function. However, A has a chi-squared distribution with one degree of freedom. Thus, the right side of Eq. (E.5) is unity.

Equation (E.5) has been used by various authors, e.g., Whiteman (59), Dreste (5), and Krohn (60). Dreste and Krohn suggest using Eq. (E.5) under the assumption "...that component-part value ranges are not large relative to the nominal value..."

Assuming that the B's are statistically independent, that the higher order terms in Eq. (E.1) can be neglected, and that Eq. (E.5) is satisfied, the estimate for the variance is obtained from Eq. (E.4) as shown in Eq. (E.6)

$$(E.6) \quad s^2 \doteq s_A^2 \doteq \sum_{r=1}^k v_r^2(m_1, \dots, m_k) s_r^2$$

The estimate for the mean is given by Eq. (E.5)

LIST OF REFERENCES¹

1. Koenig, P. Transistor reliability and Air Force requirements. Proc. of Transistor Reliability Symposium. New York University Press. 1956.
 2. Perry, D. Component reliability: what's needed? Missiles and Rockets. 5:50. July, 1959.
 3. Saltz, M. Methods for evaluating reliability growth and ultimate reliability during development of a complex system. Proc. FNSRQC. 89-98.
 4. Greene, E. Designing and building reliability into control systems. District Conference Paper, AIEE Empire Tri-District Meeting. May, 1962.
 5. Dreste, F. Circuit design concepts for high reliability. Proc. SNSRQC. 121-133.
 6. Ryerson, C. The reliability and quality control field from its inception to the present. Proc. IRE. 50:1321-1338. May, 1962.
 7. Chorafas, D. N. Statistical Processes and Reliability Engineering. D. Van Nostrand Co. Princeton, N. J. 1960.
 8. Epstein, B. Estimation from life test data. IRE TRQC. RQC-9:104-107. April, 1960.
-

¹The following abbreviations are used in this list.

- Proc. FNSRQC. Proceedings of the Fifth National Symposium on Reliability and Quality Control in Electronics (I.R.E.). 1959.
- Proc. SNSRQC. Proceedings of the Sixth National Symposium on Reliability and Quality Control in Electronics (I.R.E.). 1960
- IRE TRQC. Transactions of the Professional Group on Reliability and Quality Control (I.R.E.).
- Proc. NYUCRT. Proceedings of the New York University-Industry Conference on Reliability Theory. 1958.

9. Feyerherm, M. Basic reliability considerations in electronics. Proc.FNSRQC. 119-125.
10. Connor, J. Prediction of reliability. Proc. SNSRQC. 134-154.
11. Flehinger, B. System reliability as a function of system age; effects of intermittent component usage and periodic maintenance. Op. Res. 8:30-44. Jan.-Feb., 1960.
12. Ashcraft, W. and Hochwald, W. Design by worst-case analysis; a systematic method to approach specified reliability requirements. IRE TRQC. RQC-10:15-21. Nov., 1961.
13. Suran, J. Effects of circuit design on system reliability. IRE TRQC. RQC-10:12-18. March, 1961.
14. Patterson, M. Design techniques for upgrading the reliability of weapon systems during flight-readiness check out. IRE WESCON Conv. Rec., pt. 6. 3-9. 1958.
15. Brown, G. and Dennis, R. Electronics reliability. Missiles and Rockets. 2:172-176. Oct., 1957.
16. Drenick, R. Mathematical aspects of reliability problems. Soc. Indus. and Applied Math. J. 8:125-149. March, 1960.
17. Hellerman, L. and Racite, M. Reliability technique for electronic circuit design. IRE TRQC. IQRQC-14:9-16. Sept., 1958.
18. Davis, H. Report on a reliability program for an analog computer. Proc. FNSRQC. 79-82.
19. Luebbert, W. Dynamic failure control for military electronics. IRE TRQC. PGRQC-10:43-42. June, 1957.
20. Flehinger, B. Reliability improvement through redundancy at various switching levels. IRE National Conv. Rec., pt. 6. 137-151. 1958.
21. Von Neumann, J. Probabilistic logics and the synthesis of reliable organisms from unreliable components. Automata Studies (Annals of Mathematical Studies No. 34), 43-98. 1956.
22. Depian, L. and Grisamore, N. Reliability using redundancy concepts. IRE TRQC. RQC-9:53-60. April, 1960
23. Moore, E. and Shannon, C. Reliable circuits using less reliable relays, part I. J. of Franklin Institute. 262:191-208. Sept. 1956.

24. Moore, E. and Shannon, C. Reliable circuits using less reliable relays, part II. J. of Franklin Institute. 262:281-297. Oct., 1956.
25. Koenig, H. E. and Blackwell, W. A. Electromechanical System Theory. McGraw-Hill Book Co., New York. 1961.
26. Bazovsky, I. Reliability-Theory and Practice. Prentice-Hall, Englewood Cliffs, N. J. 1961.
27. Rosenblatt, J. On prediction of system performance from information on component performance. Proc. of Western Joint Computer Conf. 85-93. 1957.
28. Rosenblatt, J. On prediction of system behavior. IRE TRQC. RQC-9:23-28. Dec., 1960
29. Benner, A. and Meredith, B. Designing reliability into electronic circuits. Proc. N.E.C. X:137-145. 1954
30. Birnbaum, Z. Life length of materials as a stochastic process. Proc. NYUCRT. 1-6.
31. Drenick, R. The failure law of complex equipment. Soc. Indus. and Applied Math. J. 8:680-690. Dec., 1960.
32. Flehinger, B. and Lewis, P. Two-parameter lifetime distributions for reliability studies of renewal processes. Proc. NYUCRT. 7-22.
33. Barlow, R. and Hunter, L. Mathematical models for system reliability (first part). Sylvania Technologist. 13:16-31. Jan., 1960.
34. Barlow, R. and Hunter, L. Mathematical models for system reliability (second part). Sylvania Technologist. 13:55-65. April, 1960.
35. Barlow, R. and Hunter, L. Criteria for determining optimum redundancy. IRE TRQC. RQC-9:73-77. April, 1960.
36. Rohn, W. Reliability prediction for complex system. Proc. FNSRQC. 381-388.
37. Ashar, K. Probabilistic model of system operation with a varying degree of spares and service facilities. Op. Res. 8:707-718. Sept.-Oct., 1960.
38. Roberson, R. An approach to system performance prediction. J. Franklin Institute. 268:85-105. Aug., 1959.
39. Blanton, H. Reliability prediction technique for use in the design of complex systems. IRE National Conv. Rec., pt. 10. 68-79. 1957.

40. Radner, R. Limit distributions of failure time for series-parallel systems. Proc. NYUCRT. 163-186.
41. Wirth, J. Time domain models of physical systems and existence of solutions. Ph. D. Thesis. Michigan State University. 1962.
42. Caldwell, S. H. Switching Circuits and Logical Design. John Wiley and Sons, New York. 1958.
43. Lipp, J. Topology of switching elements vs. reliability. IRE TRQC. PGRQC-10:21-33. June, 1957.
44. Moskowitz, F. Analysis of redundancy networks. AIEE Trans., pt. 1. 77:627-632. Nov., 1958.
45. Weiss, G. and Kleinerman, M. On the reliability of networks. Proc. N.E.C. X:128-136. 1954.
46. Balaban, H. Some effects of redundancy on system reliability. Proc. SNSRQC. 388-402.
47. Price, H. Reliability of parallel electronic components. IRE TRQC. RQC-9:35-39. April, 1960.
48. Cramer, H. Mathematical Methods of Statistics. Princeton University Press. 1946.
49. Middleton, D. An Introduction to Statistical Communication Theory. McGraw-Hill Book Co., New York. 1960.
50. Taussky, O. and Todd, J. Generation and testing of pseudo-random numbers. Symposium on Monte Carlo Methods. John Wiley and Sons, New York. 15-28. 1954.
51. Murphy, R. Some statistical techniques in setting and obtaining tolerances. Proc. NYUCRT. 63-90.
52. Heyne, J. On an analytical design technique. IRE National Conv. Rec., pt. 6. 121-122. 1958.
53. Gray, H. Application of piecewise approximation to reliability and statistical design. IRE Proc. 47:1226-1231. July, 1959.
54. Golubjatnikov, O. Systematic approach to transistorized equipment design. Proc. FNSRQC. 21-35.
55. Xavier, M., Schneider, L. and Gottfried, P. Utilization of component part reliability information in circuit design. IRE TRQC. PGRQC-14:60-68. Sept. 1958.
56. Doob, J. L. Stochastic Processes. John Wiley and Sons, New York. 1952.

57. Feller, W. An Introduction to Probability Theory and Its Applications, Vol. 1. 2nd Ed. John Wiley and Sons, New York. 1957.
58. Bartlett, M. S. An Introduction to Stochastic Processes. Cambridge University Press. 1960
59. Whiteman, I. Reliability starts with the design. Proc. FNSRQC. 98-102.
60. Krohn, C. Reliability analysis techniques. Proc. IRE. 48:179-192. Feb. 1960.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03071 4426